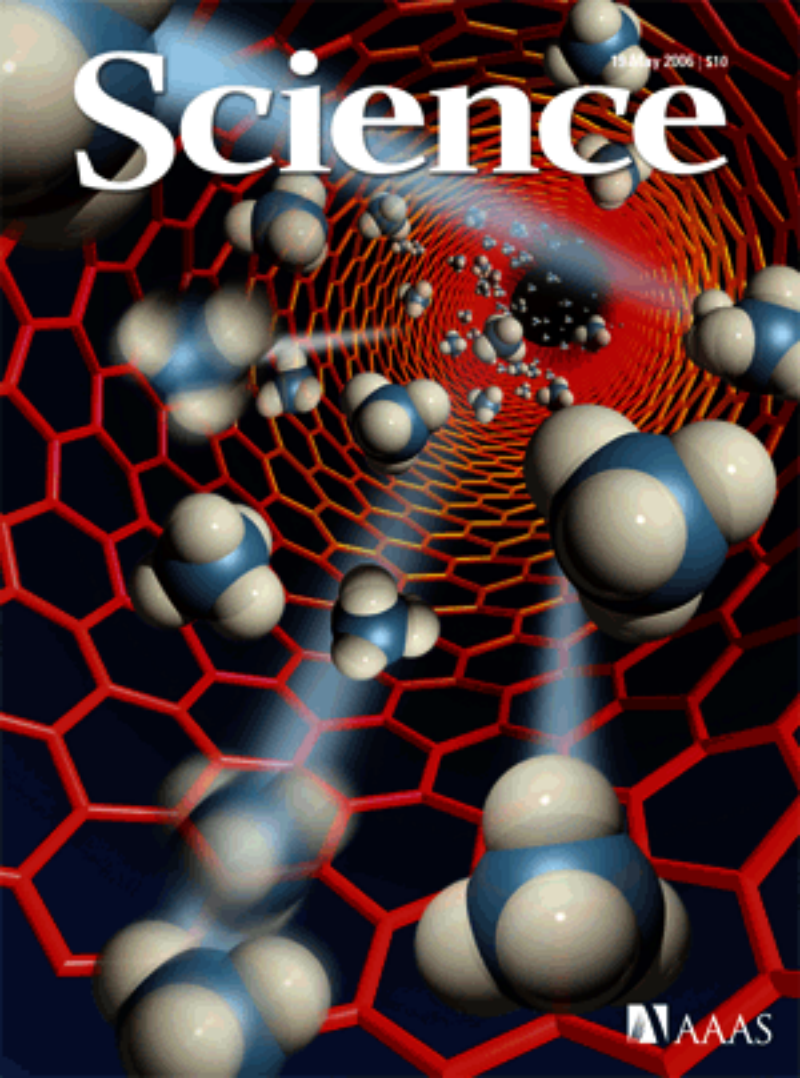
19 May 2006 | $10

Science



AAAS

EDITED BY STELLA HURTLEY AND PHIL SZUROMI

Controlling the Coming of Catkins

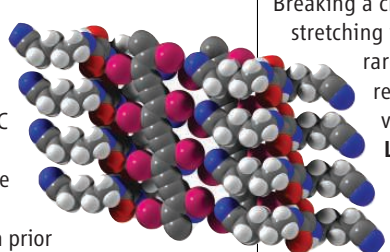
The regulation of flowering in annual plants is controlled by a pair of genes, *FT* and *CO*. Are similar genes also involved in regulating flowering time in trees? The juvenile phase in trees can last for decades before the first flower is formed. During this time, the tree is nonresponsive to environmental factors that potentially influence flowering time. **Böhlenius *et al.*** (p. 1040, published online 4 May) show that the *FT* ortholog from *Populus* trees (poplars, aspen, and cottonwoods) is a critical determinant of flowering time in trees. The poplar *FT* is also responsible for a completely different developmental process in trees, the timing of the short-day induced growth cessation and bud set that occurs in the fall.

Heavy Metal and Hard Rock

Drilling through a complete sequence of layers of the Earth's crust into the underlying pristine igneous rocks is a major goal of earth science. The thinnest crust occurs near fast-spreading mid-ocean ridges, so bore holes have targeted these regions. **Wilson *et al.*** (p. 1016, published online 20 April) drilled a 1.6-kilometer-deep bore hole through intact crust near the East Pacific Rise to reach gabbro, a layer of dark crystalline igneous rock formed from solidified magma that underlies much of the Earth's ocean floor. Determining the depth to gabbro layers confirms that magma chambers form at shallow levels in the crust at very high spreading rates; gabbros are brought up into these chambers from depth. Also, seismic bands do not correspond to compositional rock layers, implying that seismic velocities are controlled more by porosity than rock type.

Poised for Polymerization

The networks of conjugated π-orbitals in conducting polymers are stabilized either by bulky polyatomic side groups or phenyl groups incorporated within the backbone chains. **Sun *et al.*** (p. 1030; see the Perspective by **Baughman**) have prepared a polymer composed of strictly alternating C=C double and C≡C triple bonds, with only iodine atoms as side groups. The synthesis relied on prior templating of the diiododiacetylene monomer in a cocrystal with a dinitrile oxalamide host. Packing contacts between the nitriles and iodines allow the host structure to align the monomers in a column with relative spacings that changed little after polymerization. The product is a potential precursor to carbyne, a hypothesized but elusive linear allotrope of carbon.



CREDITS: (TOP TO BOTTOM) BÖHLENIUS *ET AL.*; SUN *ET AL.*

Spin Sequences

The control of coupling between spins in small structures could find use in spintronics and quantum computing. **Hirjibehedin *et al.*** (p. 1021, published online 30 March; see the Perspective by **Brune**) assembled chains of Mn atoms with a scanning tunneling microscope on a thin insulating surface (a monolayer of CuN grown on a Cu surface). They then used inelastic tunneling spectroscopy to measure spin excitation spectra as a function of chain length (up to 10 atoms) under cryogenic conditions. Comparison of the spectra with a Heisenberg model of an open spin chain with antiferromagnetic exchange coupling revealed the collective spin configurations as well as the strength of the coupling.

Selective Si–H Scission

Breaking a chemical bond by exciting its stretching vibration is an appealing idea that rarely works because the energy redistributes rapidly into other vibrational or rotational motions. **Liu *et al.*** (p. 1024; see the Perspective by **Tully**) found that H atoms adsorbed on a silicon (111) surface can be expelled as H_2 by irradiation with intense pulses of infrared light tuned to the Si–H stretching frequency. Although local heating of the surface could also cleave the Si–H bonds, the authors rule out this thermal mechanism by irradiating a mixture of adsorbed H and D atoms under the same conditions. Whereas simple heating of the surface favors D_2 over H_2 production by ~17:1, resonant excitation of the Si–H stretch reverses the selectivity to favor H_2 by a factor of 200.

Beating a Bottleneck

Knudsen diffusion occurs when the mean free path of atoms or molecules is relatively long compared to the pore or channel through which they move, so that wall collisions become more frequent than those between particles. This model holds for pores between 2 and 50 nanometers, but what happens during flow in smaller channels? **Holt *et al.*** (p. 1034, see the cover and the Perspective by **Sholl and Johnson**) fabricated membranes using double- and multiwalled carbon nanotubes to form the pores. For gases, flow rates were an order of magnitude greater than those predicted by Knudsen diffusion, and water flow rates greatly exceeded values calculated from hydrodynamics. The authors argue that the enhanced transport is caused by the smoothness of the inner nanotube surfaces, in agreement with results from computer simulations.

Evolution by Reduction?

The origins of eukaryotes remain controversial and somewhat enigmatic. **Kurland *et al.*** (p. 1011) provide a counterpoint to current models in which the eukaryotic cell is derived from structurally and genetically less complex prokaryotic cells. On the basis of genomic and

Continued on page 971

Published by AAAS

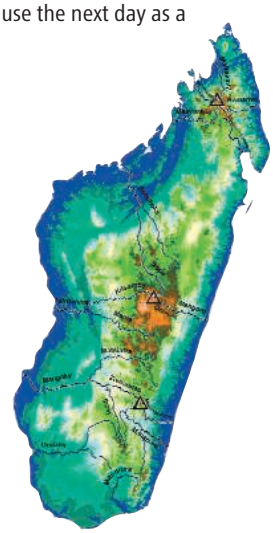
Continued from page 969

proteomic evidence, they suggest that the essence of eukaryotic cellular complexity existed in the common ancestor of eucarya, bacteria, and archaea, and that the bacteria and archaea have evolved by genome reduction driven by specialization for fast growth and cell division and/or adaptation to extreme environments.

Being Prepared

Planning ahead requires a host of cognitive skills, not the least of which is the capacity to foresee a future state of need, provocatively termed mental time travel. There is persuasive evidence that scrub jays can relocate their food caches to avoid losing them to their observant neighbors, thus preserving them for future consumption (see **Dally *et al.***, published online May 18). **Mulcahy and Call** (p. 1038; see the Perspective by **Suddendorf**) present a series of experiments that assess whether bonobos and orangutans can be coaxed to display these skills. Both species of great apes can select a suitably useful object, keep it with them overnight, and bring it back for use the next day as a tool for retrieving a food reward.

Special Speciation in Madagascar



A high percentage of the fauna and flora of Madagascar is endemic to the island, a consequence of its isolation from the African mainland since the Jurassic. Madagascar is also noted for a high degree of local endemism within the island, often to particular watersheds, a pattern that has long puzzled biogeographers. On the basis of a database of species distributions in relation to rivers and watersheds, **Wilmé *et al.*** (p. 1063) show how patterns of climate fluctuation have reinforced local isolation of populations, particularly of forest-dwelling species, to give rise to conditions suitable for speciation on a local scale.

Marshalling DNA Defenses

Cells recognize damaged DNA and initiate a complex signaling mechanism that help cells cope with the damage and initiate repair. But it is not just the enzymes required for DNA repair that undergo increased expression in response to DNA damage—other events such as progression through the cell cycle, stress responses, and metabolic pathways are also regulated. **Workman *et al.*** (p. 1054) used a systems-level approach to map such signaling pathways that respond to DNA damage. The results revealed unanticipated regulatory interactions and pave the way to when such maps may be used to predict the patient-specific effects of particular drugs.

Who Gets the Credit?

In working backward from outcomes to behavior or in strategic planning for future scenarios, one important issue is who gets the credit (and how much) for the eventual result. In the trust game, the first player has to decide how much money to invest, and the second player has to decide how much of the multiplied investment to give back. **Tomlin *et al.*** (p. 1047) have carried out a large-scale simultaneous brain imaging study and suggest that different regions of the cingulate cortex become active when what the "other" player has chosen to do is revealed, compared with situations when "I" have done the choosing.

Controlling the Synapse

Synapses in the neuromuscular junction are key components involved in control of muscle movement. **Kittel *et al.*** (p. 1051, published online 13 April; see the Perspective by **Atwood**) describe the role of *Drosophila* Bruchpilot (BRP), a coiled-coil domain protein, in the establishment and maintenance of synapses. BRP was localized to donut-shaped structures centered at the transmitter release sites (active zones) of *Drosophila* neuromuscular synapses. In mutants lacking BRP, presynaptic membranes were defective. The authors suggest that BRP is needed to form a fully functional synapse and might mediate presynaptic changes in vivo by establishing a close proximity between Ca^{2+} channels and vesicles at release sites.

CREDIT: WILMÉ *ET AL.*

Published by AAAS

Science as Smoke Screen

Stephen C. Trombulak is professor of Biology and Environmental Studies at Middlebury College in Vermont and president of the Society for Conservation Biology–North America Section.

David S. Wilcove is professor of ecology, evolutionary biology, and public affairs at Princeton University.

Timothy D. Male is senior ecologist at Environmental Defense.

THE KAUAI CREEPER, A SPARROW-SIZED SONGBIRD RESTRICTED TO THE HAWAIIAN ISLAND OF KAUAI, would seem to fit anyone's definition of an endangered species. Fewer than 1500 individuals survive in an area of only 86 km^2; its numbers are declining and it is under assault from non-native predators, pathogens, and competitors. Despite having been listed as "Critically Endangered" by the World Conservation Union, the Kauai Creeper hasn't yet earned a place on the U.S. endangered species list. In this respect, it has plenty of company. Thousands of U.S. species in grave danger of extinction have yet to be accorded protection under the Endangered Species Act of 1973 (ESA).

Fortunately, 1300 species, subspecies, and populations have been given protection under the ESA. Some, like the bald eagle, have recovered to a level that allows them to be removed from the list. Many others, although not yet out of danger, have been saved from extinction because of protection provided by the ESA. Furthermore, the U.S. Fish and Wildlife Service (FWS) has made very few errors in listing species during the 33 years it has been administering the ESA; only 10 of 1300 species (<1%) have had to be delisted because subsequent information indicated that the original decision to protect them was erroneous.

Given the enormous backlog of unprotected species in danger of extinction, one would expect Congress to expedite their protection. Instead, a bill to overhaul the ESA that passed the U.S. House of Representatives in September 2005 (H.R. 3824) would make it harder to protect endangered species, and similar "reform" legislation is now being discussed in the Senate. The pretense for the bill is to improve science, but instead H.R. 3824 would limit the use of well-tested population models for determining whether to add a species to the endangered list or for setting recovery goals. It would also add layers of time-consuming review before recovery plans could be finalized or federal agencies could act to help endangered species. Such changes will make the ESA neither scientifically sounder nor more effective.

Concerned that the scientific foundation of the ESA could be weakened by these sorts of changes, 17 scientific societies, including the Society for Conservation Biology–North America, Ecological Society of America, American Fisheries Society, Entomological Society of America, Society for Range Management, and The Wildlife Society, recently released a statement on the use (and misuse) of science in the ESA.* The statement concludes that the FWS already has effective processes in place to gather and use the best available scientific information for decision-making. However, the groups recommended the creation of an independent science advisory panel, similar to those used at the Environmental Protection Agency and elsewhere, to advise the Secretary of the Interior on issues where significant scientific uncertainty exists.

Earlier protection of rare and declining species, before they reach the brink of extinction, will greatly increase the probability that those species can be recovered. The FWS should work with the scientific community to develop clear quantitative criteria for identifying what constitutes an endangered species. Similar criteria were developed by scientists and adopted in 2001 by the World Conservation Union. The new criteria ensured consistency in determining which species should be considered imperiled.

None of this can happen unless the agencies in charge of implementing the ESA have adequate funding. The median expenditure per listed species in 2004 was only about $5500. Even this figure is somewhat deceptive because a mere 50 species (out of 1300) received 84% of all funds from the FWS and National Oceanic and Atmospheric Administration. Ultimately, too many species are dwindling for lack of attention because there isn't enough money to pursue conservation research and recovery actions in their interest. A recent study by environmentalists† recommended an increase of $68 million in the annual budget (which is probably a conservative figure).

Critics of federal regulatory policies often plead for "sound science," a cryptic rallying cry for those who really want to discourage regulation. Congress shouldn't be allowed to get away with using it as a smokescreen for eviscerating an important and successful law like the ESA. Congress did well in unanimously supporting the designation of 11 May as the first official "Endangered Species Day." However, they'll need to do more than that to show that America's commitment to the goals of the ESA is serious.

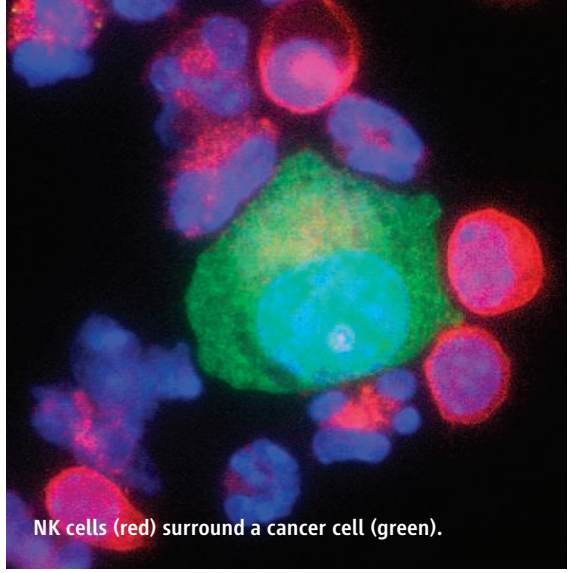
Stephen C. Trombulak, David S. Wilcove, Timothy D. Male

10.1126/science.1129508

*www.conbio.org/Sections/NAmerica/ScientificSocietiesOnUSESA.pdf. †www.defenders.org/greenbudget.pdf.

CREDIT: JACK JEFFREY PHOTOGRAPHY

Published by AAAS



NK cells (red) surround a cancer cell (green).

BIOMEDICINE

Resisting Renegade Cells

Despite the many examples involving experimental or clinical stimulation of immune responses to tumor cells, it is not yet clear to what extent the immune system might be able to combat or suppress malignancy on its own. The spontaneous remission/complete resistance (SR/CR) strain of mice is unusual in that it strongly resists challenges with high-dose inoculations of tumor cells that would otherwise be lethal. This resistance segregates as a single-locus dominant trait and correlates with significant leukocyte infiltration of the cancer.

Building on their earlier findings, Hicks *et al.* report that the infiltrate contains a variety of leukocyte subsets, including T cells, natural killer (NK) cells, neutrophils, and macrophages. Direct contact and killing of tumor cells by these immune response effectors could be measured in vitro, and resistance to both new and established cancers was conferred on wild-type mice by adoptive transfer of either bone marrow or other leukocyte fractions. Notably, SR/CR resistance was maintained even after depletion of B and T cells, revealing an innate immune component of the phenotype. The tantalizing possibility exists that characterization of this locus will improve our understanding of immune-mediated resistance to malignancy. — SJS

Proc. Natl. Acad. Sci. U.S.A. **103**, 10.1073/pnas.0602382103 (2006).

MICROBIOLOGY

Designer Pad

Legionella pneumophila, an opportunistic bacterial pathogen responsible for Legionnaire's disease, reproduces inside specialized vacuoles after phagocytosis by its host cells—either free-living protozoa or human macrophages. *Legionella*-containing vacuoles do not fuse with other endocytic vesicles but instead recruit vesicles from the early secretory pathway. They modify the vacuole membrane by using a type IV secretion system, which transports effector proteins made by the bacterium into the host cell.

Weber *et al*. examined the role of host-derived phosphoinositides (PIs) in intracellular replication and found that they are important in the anchoring of secreted bacterial effector proteins inside the vacuole. Specific effector proteins interact with a variety of host-derived PIs and, in particular, recruit PI(4) phosphate in order to attach themselves to the vacuolar membrane. Mutant bacteria lacking functional type IV secretion systems fail to modulate host cell PI metabolism and are degraded. — SMH

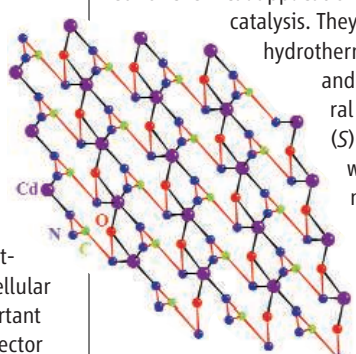
PLoS Pathog. **2**, e46 (2006).

MATERIALS SCIENCE

Twisting MOFs into Ferroelectrics

Ferroelectric materials, in which the bulk lattice exhibits a spontaneous net dipole moment, have numerous applications as memory elements, insulators, and resonators. Typical ferroelectrics are composed of inorganic salts such as $BaTiO_3$ and $LiNbO_3$, but there is interest in finding organic or organometallic alternatives.

Ye *et al*. explored the potential for ferroelectricity in a metal-organic framework (MOF) architecture, a porous motif that has been studied for chemical applications such as sorption or catalysis. They found that hydrothermal reaction of $CdCl_2$ and NaN_3 with homochiral *N*-(4-cyanobenzyl)-(*S*)-proline yields a MOF with the necessary noncentrosymmetric lattice symmetry; x-ray crystallography revealed a slight displacement of the Cd atoms in their octahedral sites. The temperature dependence of the dielectric loss suggests that the Cd-Cl bond vibration or the displacement of the proton on a tetrazoyl group (the adduct of azide-to-nitrile cycloaddition) underlies the relaxation process, and the authors estimate a dielectric constant of ~40 for this material at ~220 K. — PDS



Simplified MOF lattice depiction along the *b* axis (long red lines are benzyl groups; C-N-N and O-N units are tetrazole and proline moieties, respectively).

J. Am. Chem. Soc. **128**, 10.1021/ja060856p (2006).

APPLIED PHYSICS

Shaking Up Viscous Fluids

The transition from smooth laminar flow to chaotic turbulent flow is a problem of fundamental interest and is also of practical relevance in areas ranging from manufacturing to weather pattern formation. In Newtonian fluids such as pure water, the transition arises as a consequence of an increase in the flow rate, which in turn causes bifurcations in the flow that lead to localized flow rolls and then to chaotic or turbulent flows; in viscous fluids, these inertial instabilities are suppressed, but turbulent-like transitions have nonetheless been observed.

Schiamberg *et al.* used a parallel plate rheometer to study a series of polymer solutions in which instabilities arise from the elastic motions of individual polymer chains as they stretch and contract within a less viscous solvent. On slowly increasing the flow stress, the authors observed secondary flows: first, axially symmetric rings that formed near the outer edge of the sample; then, with rising shear stress, competing nonsymmetric rings that led to chaotic multispirals and eventually to elastic turbulence, with an accompanying factor of 13 rise in the apparent viscosity (or resistance to flow). Changing the polymer concentration induced additional flow modes, offering a rich library for theoretical development and comparison with the inertial transitions seen in Newtonian fluids. — MSL

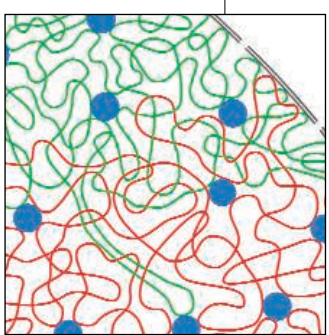
J. Fluid Mech. **554**, 191 (2006).

CREDITS (TOP TO BOTTOM): HICKS *ET AL., PROC. NATL. ACAD. SCI. U.S.A.* **103**, 10.1073/PNAS.0602382103 (2006); YE *ET AL., J. AM. CHEM. SOC.* **128**, 10.1021/JA060856P (2006)

MOLECULAR BIOLOGY

Ice Fishing

Double-stranded DNA is a wonderfully stable repository of information, as can readily be seen in the macroscopic threads of salmon sperm DNA. Compacting and condensing it into higher-order structures such as chromatin and chromosomes protects that information and allows it to fit into the nucleus. Gene expression, however, demands access to unwrapped and unwound DNA strands, which opens the door to unplanned and unwanted double-stranded breaks. These moments of vulnerability touch on a currently debated issue: the relative spatiotemporal distributions of chromosomes, with respect to each other and to transcriptionally active nuclear regions.



A close-up of intermingling between chrosome territories.

Branco and Pombo have adapted fluorescence in situ hybridization for use on ultrathin cryosections and examined how much of each chromosome territory mixes with that of the others (roughly 40% on average). They go on to show that activating expression (by applying interferon-γ to lung fibroblasts) from the MHC class II locus on chromosome 6 increases the penetration of this region into the territories of other chromosomes. Finally, the intriguing correlation between the amount of intermingling in human lymphocytes, calculated for pairs of chromosomes, and previous measurements of translocation frequencies in the same cell type highlights the importance of happenstance in rearrangements. — GJC

PLoS Biol. **4**, e138 (2006).

CHEMISTRY

Bend Origins

Chemical paradigms for multiple bonding were recently challenged by the synthesis of a chromium dimer that appeared to be held together by the interaction of 10 electrons between the Cr centers (see Nguyen *et al.*, Reports, 4 November 2005 p. 844). Before this discovery, isolable compounds were limited to bonding motifs in which eight or fewer electrons were shared between any two atoms. Orbital conformations in a quintuple bonding framework were largely expected to induce a linear geometry, but the bulky triaryl ligands capping the Cr centers adopted a bent, mutually trans configuration.

Brynda *et al.* have analyzed this geometrical conundrum using high-level quantum-mechanical calculations incorporating multiconfigurational perturbation theory. For a model compound with phenyl groups in place of the triaryl ligands, the linear conformer was energetically favored over the bent form by only 1 kcal/mol. Orbital occupancy analyses were consistent with participation of all 10 electrons in both conformers, though with repulsive antibonding contributions lowering effective bond orders to 3.69 and 3.52 for the linear and bent forms, respectively. — JSY

Angew. Chem. Int. Ed. **45**, 10.1002/anie.200600110 (2006).

Science stke AAAS

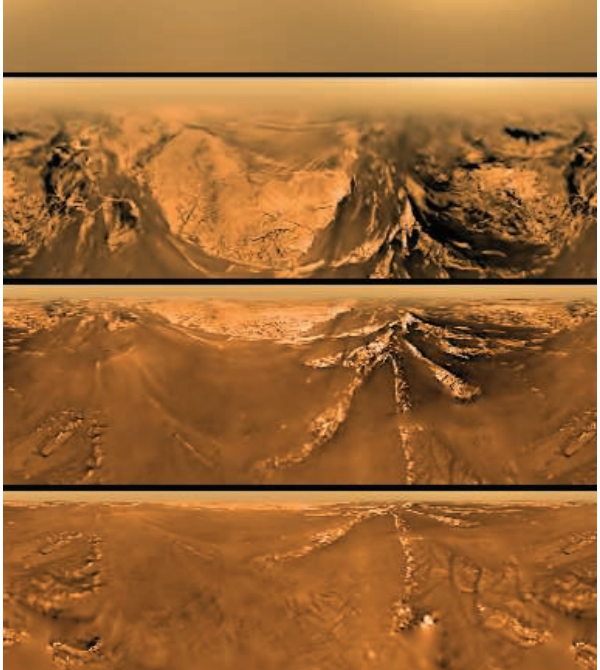
www.stke.org

<< PTEN Affects Brain Development

Mutations in the tumor suppressor *PTEN* (phosphatase and tensin homolog on chromosome ten) are associated not only with tumor development but also with several brain disorders. Intriguingly, *PTEN* mutations have been reported in individuals with autism spectrum disorders (ASD) occurring in conjunction with macrocephaly. Kwon *et al.* used mutant mice in which *Pten* was deleted in a subset of differentiated neurons in the hippocampus and cerebral cortex to investigate the effects of PTEN on brain development and behavior. The mutant mice exhibited behavior evocative of that of individuals with ASD: atypical social interactions, exaggerated responses to stressful sensory stimuli, and atypical responses in paradigms designed to assess anxiety and learning. Their brains were enlarged in the regions in which *Pten* was deleted; this was associated with hypertrophy of the cell bodies of *Pten*-negative neurons as well as increased and abnormal growth of neuronal processes. The hypertrophied neurons showed increased phosphorylation of downstream targets of Akt signaling. Thus, abnormal activation of Akt signaling in a subset of neurons appears to promote macrocephaly and behaviors that resemble some of those associated with ASD. — EMA

Neuron **50**, 377 (2006).

CREDIT: BRANCO AND POMBO, *PLOS BIOL.* **4**, E138 (2006)



IMAGES

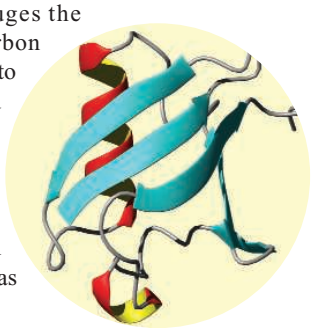
By the Light of a Coppery Moon

In January 2005, the Huygens space probe parachuted onto the surface of Saturn's moon Titan. Now you too can take the plunge, thanks to these new movies from NASA, the European Space Agency, and the University of Arizona, Tucson. The videos, the first to record the landing, condense several hours of data nabbed by the spacecraft's Descent Imager/Spectral Radiometer. In one movie, the probe dives into a thick fog and then emerges over a rugged landscape that looks like it's made of copper (above). Viewers follow Huygens all the way to its touchdown in a dry riverbed, where it nestles among pebbles and lumps of ice. The second "bells and whistles" video adds a readout of the craft's trajectory and other data. >> www.nasa.gov/mission_pages/cassini/media/cassini-20060504.html

DATABASE

Go for a Spin

Looking for nuclear magnetic resonance (NMR) spectroscopy data for ubiquitin (below), a cellular tag for worn-out molecules? You can find NMR results for ubiquitin and more than 3700 other molecules at the Biological Magnetic Resonance Data Bank from the University of Wisconsin, Madison. NMR spectroscopy gauges the nuclear spins of atoms such as carbon and hydrogen, allowing researchers to deduce molecular structures and identify compounds in chemical mixtures. The information in the database comes from the literature and researcher contributions. A new section on metabolomics houses data on small molecules that cells manufacture, such as amino acids and sugars. >> www.bmrb.wisc.edu



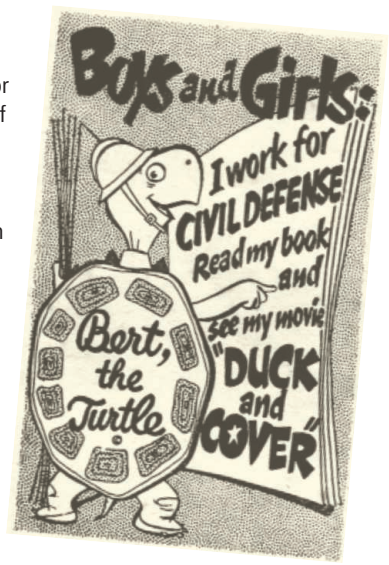
Send site suggestions to >> netwatch@aaas.org
Archive: www.sciencemag.org/netwatch

CREDITS (CLOCKWISE FROM TOP LEFT): NASA/ESA/JPL/UNIVERSITY OF ARIZONA; CONELRAD; BMRB

EXHIBIT

Ducking the Bomb

It was a time when New York City schoolchildren received dog tags so that their bodies could be identified after a nuclear attack. Mutant monsters swarmed across America's TV and movie screens, and songs like "Your Atom Bomb Heart" and "Radioactive Mama" hit the airwaves. For a cheeky history of the A-bomb's impact on popular culture, tune to CONELRAD, named for the emergency broadcasting system of the 1950s and 1960s. The Web site's personnel—a retired U.S. Air Force officer and two "civilian veterans" of the Cold War, a pop music historian and an editor—have compiled a thick dossier of rare nuclear-age memorabilia. Read the history of the famous civil-defense film *Duck and Cover* (right), or spin selections from the 1961 instructional record "If the Bomb Falls," whose advice for stocking a fallout shelter included packing plenty of tranquilizers. >> www.conelrad.com



TOOLS

There Goes the (Genetic) Neighborhood

Researchers use linkage analysis to map disease-causing genes, but calculations that involve complicated human pedigrees often stump the average computer. Superlink Online from the Technion–Israel Institute of Technology in Haifa overcomes this limitation by farming out the number-crunching to a network of some 2700 PCs, which tackle the calculations during their spare time. Described online in the *American Journal of Human Genetics* this month, the site computes the likelihood that genes lie within a particular chromosome neighborhood and can handle larger pedigrees than other linkage software. After obtaining a free password, users feed their own data into the program. They can also add their machines to the network. >> bioinfo.cs.technion.ac.il/superlink-online

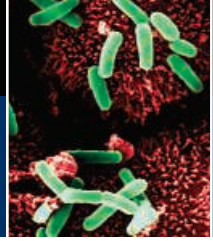
WEB LOG

Invasion Chronicles

An outbreak of pine shoot beetles (*Tomicus piniperda*) has prompted the U.S. Department of Agriculture to restrict the export of bark chips and other forest products from Massachusetts, Connecticut, and Rhode Island. Meanwhile, farmers in southwestern Puerto Rico are angry because the government has failed to control hungry mobs of Asian and African monkeys, descendants of escapees from a medical lab, that are pillaging their fields. For more news about wayward organisms and efforts to control them, check the Invasive Species Weblog from ecologist Jennifer Forman Orth of the University of Massachusetts, Boston. Orth gleans the postings from media stories, government and university announcements, reports by professional societies, and other sources from around the world. >> invasivespecies.blogspot.com

Published by AAAS

New twists in the hobbit saga
983



Taking the measure of *Clostridium*
986

KOREAN CLONING SCANDAL

Prosecutors Allege Elaborate Deception and Missing Funds

SEOUL—Once-famed, now-disgraced stem cell pioneer Woo Suk Hwang was indicted on 12 May on charges of fraud, embezzlement, and violations of a bioethics law. Five other members of his team have also been indicted, three on fraud charges, one on a bioethics law violation, and one for destroying evidence and obstructing business operations. Hwang claims that he has been falsely accused on several points, according to Geon Haeng Lee, one of Hwang's seven lawyers.

Hwang, formerly a professor at Seoul National University (SNU), had claimed in a 2004 *Science* paper (12 March 2004, p. 1669) to have made a breakthrough in so-called therapeutic cloning by creating a stem cell line from a cloned human blastocyst. He followed that up a year later with a second *Science* paper claiming to have created 11 stem cell lines derived from tissue contributed by patients suffering from spinal cord injury, diabetes, or an immune disorder (17 June 2005, p. 1777). Together, these papers seemed to pave the way toward creating replacement cells and tissues for these and other diseases that would be genetically matched to individual patients. Hwang was feted by scientists around the world and became a national hero in South Korea, which hoped to ride his achievements to worldwide prominence in stem cell research.

The claims started unraveling last fall. First, bioethical lapses in collecting oocytes were alleged, then problems with manipulated photos and other supporting data were identified (*Science*, 23 December 2005, p. 1886). In January 2006, SNU announced that an investigating committee had concluded that no cloned stem cell lines existed. Hwang and his co-authors retracted both papers, and Seoul public prosecutors launched an investigation (*Science*, 6 January, p. 22).

The prosecutors' conclusions are documented in a 150-page report that fills in some of the remaining holes in the Hwang saga. According to the prosecutors, Hwang and his team apparently believed that the "number 1" stem cell line that formed the basis for the 2004 *Science* paper was truly derived from a cloned blastocyst. Two separate investigations by SNU, however, concluded that the blastocyst most likely resulted from parthenogenesis, a form of asexual reproduction. The prosecutors' report leaves it up to academics to sort out whether the blastocyst was the result of cloning or parthenogenesis.

Sweeping charges. In Gyu Lee of the Korean public prosecutor's office released a long-awaited report on the cloning scandal, indicting Woo Suk Hwang and five others on charges including fraud and embezzlement.

However, the report says Hwang's team did not keep proper records and did not have evidence to support any scientific claims about stem cell line number 1. So, the prosecutors allege, Hwang ordered associates Jong Hyuk Park and Sun Jong Kim to fabricate photos, DNA test results, and other supporting data for the 2004 *Science* paper.

For the June 2005 paper claiming the creation of 11 patient-specific cell lines, the report says that Kim, a member of the team from MizMedi Hospital in Seoul, was in charge of deriving stem cells from cloned blastocysts that had been created at the SNU lab. He was unable to do so. But, the report says, feeling pressure to perform and wanting to make a name for himself, he took fertilized stem cells from MizMedi's collection and mixed them with material from Hwang's lab. He reportedly told other researchers that light was "not good for the cells" and did most of the work in semidarkness. Prosecutors concluded that no one else in the lab, including Hwang, realized what had been done until suspicions were raised after the paper was published, when DNA fingerprinting tests in December 2005 showed that the customized stem cell lines were identical to the fertilized stem cells from MizMedi.

The report alleges that Kim created two lines, and Hwang, believing they were real, ordered him to fabricate data to make it look as though they had made 11. Kim was indicted for obstructing research work at SNU, as well as for destroying evidence. The prosecutors allege that, in addition to deleting related computer files from his laptop and computers at MizMedi, Kim told MizMedi researchers to hide the fact that he was removing stem cells from its labs.

Although Kim allegedly deceived Hwang, the prosecutors say that Hwang was ultimately responsible for ordering subordinates to fabricate data. The prosecutors did not file any charges against Hwang for publishing fraudulent research reports, however, saying it would be a complicated procedure that would have to involve *Science*.

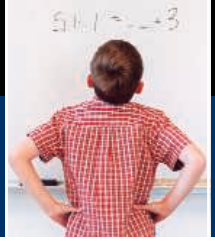
The prosecutors confirmed earlier reports that Hwang had used many more oocytes than the several hundred he acknowledged, collecting 2236 eggs from 122 women, 71 of whom were compensated. Paying for oocytes continued even after a bioethics law banned the practice in January 2005, the prosecutors' report states.

Meanwhile, in addition to research misconduct, the prosecutors claim Hwang misappropriated $2.99 million in state funds and private donations. Their report outlines an elaborate scheme in which Hwang withdrew large amounts of cash and carried it in bags to other banks to avoid a paper trail of bank transfers. The prosecutors say he had 63 accounts under different names, including those of junior researchers and relatives. To cover up some of the alleged embezzlement, he wrote false tax statements claiming to have bought pigs and cows for research purposes. Hwang faces up to 3 years in prison for violating the bioethics law and up to 10 years for the misuse of state funds.

The prosecutors also indicted two of Hwang's colleagues at SNU, professors Byeong Chun Lee and Sung Keun Kang, for ▶

CREDIT: LEE JAE-WON/REUTERS

Published by AAAS



Toward a truce on math curricula
988

Prospects for ALMA looking up
990

China's groundbreaking tokamak
992

fraud. The report says the two provided false evidence in order to receive government grants and then misappropriated the money. SNU has begun taking steps to fire the two professors.

Sang Sik Chang, head of the Hanna Women's Clinic in Seoul, which provided Hwang with eggs in 2005, was charged with violations of the bioethics law in connection with egg procurement. Hyun Soo Yoon, a professor of medicine at Hanyang University in Seoul, was indicted for creating false receipts and embezzling research funds approved for a joint research project to create stem cells at MizMedi.

Sung Il Roh, director of MizMedi, who also gave oocytes to Hwang, was not indicted; prosecutors say Roh did not pay for any oocytes after the bioethics law went into effect. Shin Yong Moon, a stem cell specialist at SNU who was co–lead author with Hwang on the 2004 *Science* paper, was cleared of wrongdoing by the prosecutors.

Hwang's lawyer, Lee, says Hwang maintains that he did not order junior researchers to fabricate data for the 2004 article and that he believed a member of his team had created the number 1 stem cell line from a blastocyst resulting from somatic cell nuclear transfer, not parthenogenesis. "Prosecutors based their conclusion on testimonies from Jong Hyuk Park and Sun Jong Kim and did not take into consideration Hwang's statements that he did not order them to fabricate data," Lee says.

Hwang's lawyer also denied that Hwang embezzled funds, saying that the scientist had made huge profits from lectures and publications, which amounted to about $840,000. That money was put into the same bank accounts as his grants, but items such as his wife's car were bought with those private earnings, he contends. He says Hwang's lawyers will fight the charges in court. The first trial is scheduled for 20 June.

Meanwhile, the South Korean government says that it will try to retrieve the grant money given to Hwang and his lab at SNU. The Ministry of Science and Technology says, however, that about $3.2 million has already been spent on design and construction of a new research facility that was being built adjacent to the College of Veterinary Medicine; those funds will be considered losses. SNU has not yet decided what to do with the unfinished building.

Hwang's supporters continue to urge Hwang to restart his research and the South Korean government to acquire a patent on the first stem cell line. "Hwang may have rushed to publish the 2005 article, but he should be acknowledged for creating the first stem cell line and cloning Snuppy" the dog, one supporter says. "We have to obtain a patent for the country's sake, not Hwang's."

Last weekend, hundreds of Hwang's supporters gathered in front of the prosecutors' office, protesting Hwang's indictment. Police sealed off access to rooftops of nearby buildings to prevent suicide, as has been attempted in the past. Before the indictment, the Venerable Seol, a Buddhist monk, announced on 8 May that three individuals had pledged to contribute $65 million to help Hwang, a fellow Buddhist, restart his research. After the prosecution's announcement on 12 May, several monks began a 24-hour relay bowing ritual next to Jogye Temple in central Seoul in support of Hwang.

Looking beyond individual culpability, senior prosecutor In Gyu Lee said at a press briefing that he placed partial blame for the scandal on "the strict Korean lab culture," which leaves junior researchers powerless to refuse unethical demands by lab heads. He added that although the scandal demonstrated that "a lot of scientists lacked ethics," he also noted that the fraud had damaged many junior researchers and collaborators who had no idea of what Hwang and his close associates were up to.

South Korea's research community seems to be taking the lesson to heart, says Kye Seong Kim, a stem cell researcher at Hanyang University College of Medicine. He believes universities will now set up offices of research integrity. "That's one good thing that might come out of this tragedy," he says. Others think reforms must go further. Duck Hwan Lee, a chemistry professor at Sogang University in Seoul, places partial blame on the government for pouring so much money into Hwang's project without sufficient information. "[The government] should create a system that enables more transparent research funding. Scientists should be able to compete for grants fairly instead of relying on lobbying or personal ties," he says.

–D. YVETTE WOHN AND DENNIS NORMILE

D. Yvette Wohn is a reporter in Seoul.

CREDITS (TOP ROW, LEFT): KIM KYUNG HOON/REUTERS; (ALL OTHER IMAGES) YONHAPNEWS AGENCY

Published by AAAS

U.S. MATHEMATICS EDUCATION

Well-Balanced Panel to Tackle Algebra Reform

Calling himself an "honest broker," former University of Texas president and chemist Larry Faulkner has been named to chair a new presidentially appointed panel that will tackle the long-running debate over reforming U.S. mathematics education.

The 17-member National Mathematics Advisory Panel is part of a proposed $250 million mathematics initiative by the Bush Administration.* The Math Now initiative, aimed at giving elementary school students a strong foundation in math and boosting the abilities of middle school students who have fallen behind (*Science*, 10 February, p. 762), puts special emphasis on algebra as the key to educational success. "The president wants the best advice on promoting student readiness for algebra and higher-level courses," says Faulkner, who now heads the $1.6 billion Houston Endowment, a private philanthropy. "Algebra is a tremendously important gateway course, but our success rates are not very good."

Faulkner jokes that he was chosen "as someone with credentials in education and with the ability to massage egos." The panel, which will begin meeting next week, includes several prominent players in the ongoing debate about what teachers and students need to know and whether those needs are met by the recent curricular reforms.

Math mediator. Larry Faulkner hopes to reconcile the various views of panelists.

The two professional mathematicians on the panel—Harvard University's Wilfried Schmid and Hung-His Wu of the University of California, Berkeley—have been vocal critics of those reforms and have argued for more rigorous instruction on basic skills. Panelist Francis "Skip" Fennell is president of the National Council of Teachers of Mathematics, the nation's leading math education organization, which has championed many of those reforms, as has math educator Deborah Loewenberg Ball of the University of Michigan, Ann Arbor. But Ball and Schmid are also members of a group that has pushed to find common ground between the reformers and their critics (see p. 988). The panel's vice chair is Camilla Benbow, an educational psychologist at Vanderbilt University in Nashville, Tennessee, who co-directs a longitudinal study of gifted math students.

Education Secretary Margaret Spellings says she hopes the panel's initial recommendations, due to her in January 2007, will help U.S. teachers "know what's most effective in the classroom." The commission also has the authority to order research on related topics before submitting its final report in February 2008. Although Faulkner doesn't rule out that possibility, he says "I think quite a lot of work has already been done."

–JEFFREY MERVIS

* www.ed.gov/about/bdscomm/list/mathpanel/index.html

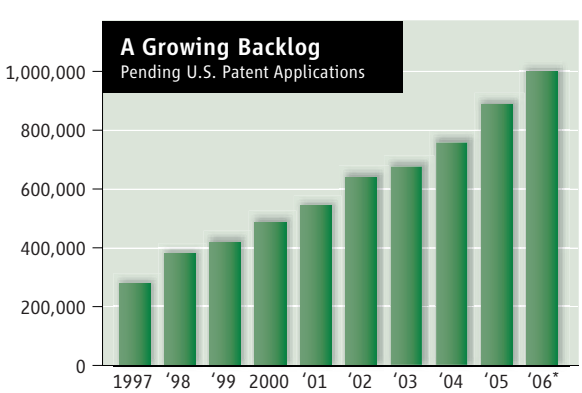
U.S. PATENT POLICY

PTO Wants to Tap Experts to Help Patent Examiners

Think someone's trying to patent an old idea? The U.S. Patent and Trademark Office (PTO) may want you to chime in.

The patent office is weighing an online pilot project to solicit public input on patent applications. Speaking last week at an open forum, officials said that tapping into the expertise of outside scientists, lawyers, and laypeople would improve the quality of patents—and might also reduce a backlog that this month topped 1 million applications. "Instead of one examiner, what if you have thousands of examiners reading an application?" says Beth Simone Noveck of New York University Law School, who is an independent advocate of the idea.

The peer initiative focuses on so-called prior art, the scientific papers and previous patents that could render claims invalid. Although applicants often flood PTO with supporting material, PTO's 4500 examiners are prohibited from consulting with outsiders about its relevance. (The law does allow outsiders to pay $180 to submit up to 10 pieces of prior art, but comments are barred to avoid the appearance of meddling.) IBM is a firm supporter of the pilot system, and PTO officials hint that software and microchip patents will be one area of focus. Former examiner Leon Radomsky says outside experts would "definitely help" those areas given the dearth of outside prior-art resources, although supporters feel that the pilot could also benefit biotechnology and the chemical sector.



Although the pilot is tentatively set to begin in December, details remain sketchy. The idea is for volunteers to be alerted about new patent applications—applications become public after 18 months—and invited to submit prior art. The community would then rank each other's suggestions, a la Amazon.com and the geek-news site Slashdot. Theoretically, says PTO official Jay Lucas, the process would generate a list of, say, 10 pieces of prior art that the examiner would do well to consult. Outsiders might also help examiners with another element of their job, namely, ruling on the tricky question of whether a proposed invention is obvious.

Some observers worry that the system will simply add to an already heavy workload for examiners. Others speculate that a competitor, assuming that an applicant would be awarded a patent, might try to game the system by not introducing some prior art until it could be used for maximum leverage as part of a later challenge to the patent. And some think PTO's problems lie elsewhere. Former patent examiner Charles Wieland III, an attorney with Buchanan Ingersoll PC, says PTO should "just let examiners develop their expertise." Inexperience is the "real problem" at PTO, he adds.

A decision on launching the project is expected this summer. **–ELI KINTISCH**

CREDITS (TOP TO BOTTOM): UNIVERSITY OF TEXAS, AUSTIN; (SOURCE) U.S. PATENT AND TRADEMARK OFFICE

Published by AAAS

PALEOANTHROPOLOGY

How the Hobbit Shrugged: Tiny Hominid's Story Takes New Turn

SAN JUAN, PUERTO RICO—The strangest ancient humans may be Indonesia's "hobbits," the 1-meter-tall people who made stone tools and hunted dwarf elephants 18,000 years ago. When announced 2 years ago, the fossils from the island of Flores seemed almost too bizarre for fiction. Now, close-up looks at some of the bones have given the hobbits' saga even more odd twists.

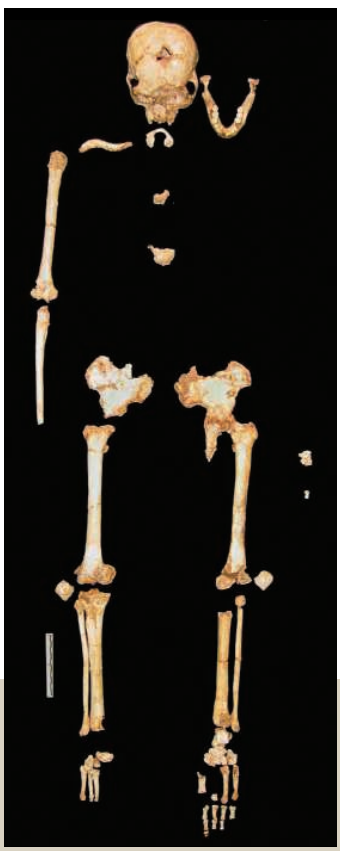
At a recent meeting here,* two anatomists presented analyses suggesting that the original hobbit skeleton may not be female, as first described, and that its shoulders differ from those of modern people and hark back to an ancient human ancestor, *Homo erectus*. That detail and others bolster the notion that an *H. erectus* population on the island evolved into the dwarf form of *H. floresiensis*, anatomist Susan Larson of Stony Brook University in New York said in her talk at the meeting.

Other researchers' opinions about almost every aspect of the hobbits, however, continue to run the gamut. Many are impressed with Larson's analysis. "I support Larson's observations … [and see] evidence of a faint phylogenetic signal" connecting the finds with *H. erectus*, says paleoanthropologist Russell Ciochon of the University of Iowa in Iowa City, who calls the skeleton from Flores "a very important link to our past." But a few researchers still find the whole tale too tall to swallow. In a Technical Comment published online this week by *Science*, paleoanthropologist Robert D. Martin of the Field Museum in Chicago, Illinois, and colleagues argue that the single skull is that of a modern human suffering from microcephaly (see sidebar). And even some researchers who are reasonably convinced that the fossils do not represent diseased modern people caution that the sample size for the shoulder bones is one. "It's always nicer to have more than one individual" to hang a hypothesis on, says Eric Delson of Lehman College, City University of New York.

At the meeting, a packed room listened intently as Larson described her work on the upper arm bone, or humerus, of the original skeleton, labeled LB1 as the first human from Liang Bua cave. The LB1 humerus is peculiar—or, rather, it lacks a peculiarity shared by living people.

In modern humans, the top or head of the humerus is twisted with respect to the elbow joint by about 145 to 165 degrees. As a result, when you stand straight, the insides of your elbows face slightly forward, allowing you to bend your elbows and work with your hands in front of your body.

But in *H. floresiensis*, the humerus appeared only slightly twisted. Last fall, Michael Morwood of the University of New England in Armidale, Australia, co-discoverer of the Flores bones, asked Larson, known for her work on the upper arm, how this could work in a toolmaking ▶



Mini-me. Details of the *Homo floresiensis* skeleton suggest that it may be descended from *H. erectus.*

CREDIT: SUSAN LARSON/STONY BROOK UNIVERSITY

* Paleoanthropology Society, 24–26 April.

But Is It Pathological?

Even as some researchers draw inferences about the ancestry of *Homo floresiensis* (see main text), others remain convinced that the bizarre bones from the Indonesian island of Flores are nothing more than diseased modern humans. In a Technical Comment published online by *Science* this week (www.sciencemag.org/cgi/content/full/312/5776/999b), paleoanthropologist Robert D. Martin of the Field Museum of Natural History in Chicago, Illinois, and colleagues make that case.

Martin gathered scaling data on the brains and bodies of other mammals, including data on the proportions of elephants as they evolved into dwarf forms on islands. Using several possible scaling models, he argues that shrinking a *H. erectus* brain to roughly the size of the Liang Bua skull would yield a body size no greater than 11 kilograms—the size of a small monkey.

If the Liang Bua bones aren't a new species of human, what are they? Martin argues that the single tiny skull may be a modern human with microcephaly, or a pathologically small head. A previous *Science* paper by Dean Falk of Florida State University in Tallahassee and her colleagues argued that the Liang Bua skull did not show the extreme pathology seen in a microcephalic brain. But Martin counters that some microcephalic brains exhibit much less pathology, including one from a 32-year-old woman reported to have had the body size of a 12-year-old child. "I'm not saying I'm 100% certain it's microcephaly," says Martin. "I'm saying that that brain size is simply too small" to be normal.

> "I'm not saying I'm 100% certain it's microcephaly … **[but the] brain size is simply too small" to be normal.**
>
> —R. D. Martin, Field Museum

Jean-Jacques Hublin of the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, who has seen the original specimens, finds the scaling arguments "quite convincing." But Martin's arguments are provoking a sharp response. Falk calls Martin's claims "unsubstantiated assertions" and adds that her team is surveying microcephalics to learn more. And bones from several small individuals have now been recovered from Flores, notes William Jungers of Stony Brook University in New York. He says that Martin's explanation implies that the island was home to "a village of microcephalic idiots." He adds that "there are precious few 'scaling laws' out there" and that examples of unusual scaling are not unexpected.

Paleoanthropologist Ralph Holloway of Columbia University, who is also studying microcephalic brains, says that so far he sees some differences between the Liang Bua skull and what's called primary microcephaly. But he warns that it will take a substantial survey to be sure. "I am coming around to believing that it isn't primary microcephaly," he says. But "I certainly would not rule out pathology just yet."

–E.C.

Published by AAAS

hominid. "I told him I didn't know," says Larson. "It *wouldn't* work."

So at the invitation of Morwood and Tony Djubiantono of the Indonesian Centre for Archaeology in Jakarta, Larson flew to Jakarta last fall to study the bones with her Stony Brook colleague William Jungers, who was to work on the lower limbs. The pair are among the handful of researchers who have studied the original specimens.

Larson found that the LB1 humeral head was in fact rotated only about 110 degrees. (No rotation would be expressed as 90 degrees.) Curious, she examined LB1's broken collarbone plus a shoulder blade from another individual.

Larson concluded that the upper arm and shoulder were oriented slightly differently in *H. floresiensis* than in living people. The shoulder blade was shrugged slightly forward, changing its articulation with the humerus and allowing the small humans to bend their elbows and work with their hands as we do. This slightly hunched posture would not have hampered the little people, except when it came to making long overhand throws: They would have been bad baseball pitchers, says Larson.

When Larson looked at other human fossils for comparison, she found another surprise: The only *H. erectus* skeleton known, the 1.55-million-year-old "Nariokotome boy" from Kenya, also has a relatively untwisted humerus, a feature not previously noted. Larson concluded that the evolution of the modern shoulder was a two-stage process and that *H. erectus* and *H. floresiensis* preserved the first step.

H. erectus expert G. Philip Rightmire of Binghamton University in New York, who works on fossils from Dmanisi, Georgia, supports this view. Larson's and Jungers's analyses "make it clearer and clearer that *Homo floresiensis* is not some sort of dwarf modern human. This is a different species from us," he says.

In a separate talk, Jungers reported more unexpected findings. He was able to reconstruct the pelvis, which had been broken when the bones were moved to a competing lab in Indonesia (*Science*, 25 March 2005, p. 1848). Although previous publications had described the pelvis as similar to those of the much more primitive australopithecines, Jungers found that the orientation of the pelvic blades is modern. The observation adds weight to the notion that hobbits had *H. erectus*, rather than australopithecine, ancestry.

The skeleton was first described as female, although the competing Indonesian-Australian team described it as male in press accounts. Now Jungers says he is "agnostic" about its sex. He notes that limb bones from other individuals from Liang Bua are even smaller—"they make LB1 look like the Hulk," he says—raising the possibility that males and females differed in size, with LB1 in the role of big male.

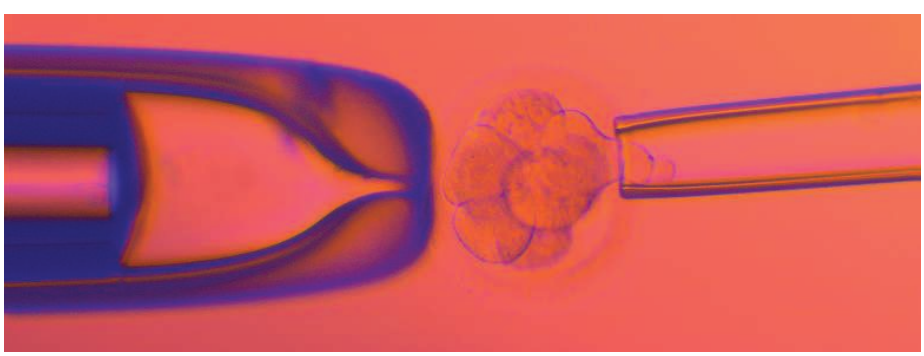
More surprises are still to come. Jungers said in his talk that LB1 includes an essentially complete foot, something not identified previously, and hinted that the foot is extremely large. Indonesia's hobbits, like J. R. R. Tolkien's fictional creatures, may have trekked about on big hairy feet.

–ELIZABETH CULOTTA

GENETIC TESTING

U.K. Embryos May Be Screened for Cancer Risk

CAMBRIDGE, U.K.—In vitro fertilization patients will be able to use genetic testing to avoid having children with mutations in genes such as *BRCA1* and *BRCA2* that raise cancer risks, the U.K. Human Fertilisation and Embryology Authority (HFEA) ruled last week. The decision, which follows a public consultation, breaks new ground because it permits screening for genes that are worrisome but not necessarily lethal or likely to produce trauma in childhood. The medical community is generally supportive, but critics are concerned that the decision could lead to screening for less risky traits in the future.



New frontier. Fertility clinics will be allowed to sample embryos before implantation for mutations in genes such as *BRCA1* and *BRCA2* and reject them.

CREDIT: PASCAL GOETGHELUCK/PHOTO RESEARCHERS INC.

Ten clinics in the United Kingdom are currently licensed to carry out preimplantation genetic diagnosis (PGD), in which one or two cells are removed from the embryo at the eight-cell stage and tested for lethal genetic conditions such as cystic fibrosis or Huntington's disease. HFEA chair Suzi Leather said on 10 May in a prepared statement that the authority's decision is "not about opening the door to wholesale genetic testing." Rather, genetic tests would be available to the minority of people with a clear history of cancer in the family. HFEA will consider applications for testing on a case-by-case basis, she says, considering factors such as family medical history and whether the condition is treatable.

Like many others in the medical community, Simon Fishel, managing director of CARE Nottingham, a U.K. clinic licensed to perform PGD, described the decision as "ethically sound." He predicts that only a very small proportion of clients will elect to use the tests. Cost will also limit take-up: Depending on how much the government contributes, patients could be left with a bill of $10,000.

But for some, the U.K. decision raises troubling questions. "I'm not entirely comfortable because of the concerns about the whole spectrum, from very severe diseases to what are essentially traits," says Francis Collins, director of the U.S. National Human Genome Research Institute in Bethesda, Maryland. "There is no bright line along that spectrum." What is most worrying, he says, is that embryo screening is not regulated in the United States, and no one is sure how widespread testing is.

Some U.K. lobby groups and disability campaigners oppose the policy outright, however, saying it smacks of eugenics. "We are concerned that people are eliminating embryos, whether they have cancer or not," says Josephine Quintavalle of the U.K. lobby group Comment on Reproductive Ethics. Quintavalle argues that research efforts should be concentrated on cancer cures, not destroying affected embryos. "We are concerned that people will view PGD as a cure for cancer," she says.

–LAURA BLACKBURN

With reporting by Jocelyn Kaiser in Washington, D.C.

Published by AAAS

HUMAN EVOLUTION

Genomes Throw Kinks in Timing Of Chimp-Human Split

A new genomic analysis has added a provocative twist to the history of humans. After comparing the genomes of five primate species, researchers have concluded that the ancestors of chimps and humans went their separate ways about 6 million years ago—at least a million years later than fossils suggest. But that's not even the most controversial claim: Early hominids interbred with their chimp cousins, says David Reich, a geneticist at Harvard Medical School in Boston. This hybridization helped make the human genome a mosaic of DNA with varying degrees of similarity to the chimp genome, he and his colleagues report in a paper published online on 17 May by *Nature*.

Human roots. New DNA studies challenge the hominid status of the 7-million-year-old Toumaï fossil (*bottom*) by suggesting that humans (*top*) and chimps (*middle*) diverged much more recently.

Researchers are impressed by the huge amount of data Reich, Nick Patterson of the Broad Institute in Cambridge, Massachusetts, and their colleagues incorporated into their study. "The paper showed that the comparative genomic approach is very powerful," says geneticist Hideki Innan of the University of Texas Health Science Center in Houston. But some, particularly paleontologists whose fossils suddenly might become too old to be hominids, are more critical. Martin Pickford of the Collège de France in Paris predicts that the work will be "of passing significance."

For decades, anthropologists have argued about the timing of the chimp-human split, with estimates ranging from 10 million to 5 million years ago. The oldest fossil put forth as a human ancestor is a spectacular skull unearthed in Chad in 2002 nicknamed Toumaï. It dates back 7 million years, says co-discoverer Michel Brunet of the University of Poitiers, France. Two other hominid species were alive in Kenya and Ethiopia 5.8 million to 6 million years ago, according to other fossils.

This fossil record doesn't neatly fit with the new findings by Reich's team. They matched up DNA sequences from the human, chimp, orangutan, macaque, and gorilla genomes and documented the differences. Having DNA from the orangutan, and from an even less related species, the macaque, allowed the group to confirm that mutations accumulated at about the same rate in different lineages of apes and humans. This meant that the number of differences in each lineage could be compared directly and were reliable for calculating how long the branches between apes and humans on the tree should be.

The sequence comparisons provided relative "genetic" ages of the five species, and based on the ages of fossils of the ancestors of orangutans and macaques, the investigators concluded that the human lineage split from chimps no more than 6.3 million years ago and perhaps even more recently than 5.4 million years ago. That timing roughly agrees with another genetic analysis, reported in December 2005, by Blair Hedges, an evolutionary biologist at Pennsylvania State University in State College. "Together, they make a strong argument against the claims of older divergence times by paleontologists and other molecular evolutionists," says Hedges.

Brunet counters that it's too early to rewrite human history based on the DNA data. "Their explanation is just a hypothesis, while Toumaï is a true fossil," he says. Also, the difference between the dates from the molecular analyses and the age of the Chad fossil may not be significant. "There are broad confidence limits on genetic data," says Montgomery Slatkin, a population geneticist at the University of California, Berkeley.

But no matter when hominid speciation occurred, the genetic analysis revealed that the transition wasn't very smooth. By comparing discrete sections of the primate genomes, Reich's team was able to calculate at least a 4-million-year difference in the ages of the oldest and youngest parts of the human genome. The X chromosome's age was most surprising. Chimp and human X chromosomes are much more similar than are the rest of their chromosomes, says Reich. Based on this congruency, he and his colleagues calculate that the X chromosomes became species-specific 1.2 million ▶

CREDITS (TOP TO BOTTOM): PHOTOS.COM; DORLING KINDERSLEY/GETTY IMAGES; MPFT

Exile for Export Rule Change

Under pressure from researchers, the U.S. Commerce Department has retreated from new export-control rules that would have made it harder for nationals from some countries to do research in the United States. One year ago, the government proposed new rules on safeguarding sensitive technologies, one of which would have required schools to obtain export licenses before employing foreigners including Indians, Chinese, and Russians in certain projects. Universities argued that the rules were so onerous that they'd hinder research (*Science*, 13 May 2005, p. 938).

Commerce now wants "to step back … and consider more broadly how best to balance national security with openness in research," says Under Secretary of Commerce for Industry and Security David McCormick. Commerce is forming a committee to review the issue and report back within a year. Academics hope any new policies will address their concerns.

–YUDHIJIT BHATTACHARJEE

Regrets Only

A going-away party for the director of the National Cancer Institute (NCI) has been postponed after questions of propriety arose. Andrew von Eschenbach, who is also acting chief of the Food and Drug Administration (FDA) and has been nominated to head the agency, was to be the subject of a 17 May reception and roast until *The Cancer Letter* questioned the invitation's statement that "gift contributions [are] also welcome." Federal ethics rules bar gift solicitations for a superior; the National Institutes of Health makes an exception if the official has resigned, but von Eschenbach hasn't yet and also regulates NCI clinical trials at FDA. NCI says the event has been postponed, and "there will not be a gift."

–JOCELYN KAISER

Canada on Kyoto: What a Gas

Two weeks after Canada's new Conservative government terminated a package of programs designed to reduce greenhouse gas emissions, Canadian Environment Minister Rona Ambrose calls Canada's Kyoto Protocol commitments "unachievable."

In a formal submission to the United Nations last week, the government explained that energy-exporting countries such as Canada "provide other countries with opportunities to switch to cleaner sources of fuel." Ambrose plans to unveil new emissions controls this fall, but activists say Ottawa is abdicating its responsibility.

–PAUL WEBSTER

Published by AAAS

years after the rest of the genomes.

To explain this oddity, Reich proposes that after evolving their separate ways for an unknown length of time, the earliest hominids and chimps hybridized. To be fertile, the hybrids had to have compatible X chromosomes, and thus there was intense selection to weed out any differences on that chromosome. Only after hybridization ceased did the X chromosome evolve into two different ones again.

Innan's analysis of just human and chimp DNA, published earlier this month in *Molecular Biology and Evolution*, supports the idea of hybridization between chimp and human ancestors. Still, Reich theory's is getting a tough reception. "I don't buy these hybrids," says Harvard anthropologist David Pilbeam, arguing that the ancestors of hominid and chimp were too different, morphologically and developmentally, to produce fertile offspring.

As more primate genomes are sequenced, the history of the X chromosome should become clearer, says Reich. Whether chimp ancestors interbred with human ancestors or not, notes Svante Pääbo of the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, comparative genomics "tells us … things that paleontology can't."

–ELIZABETH PENNISI

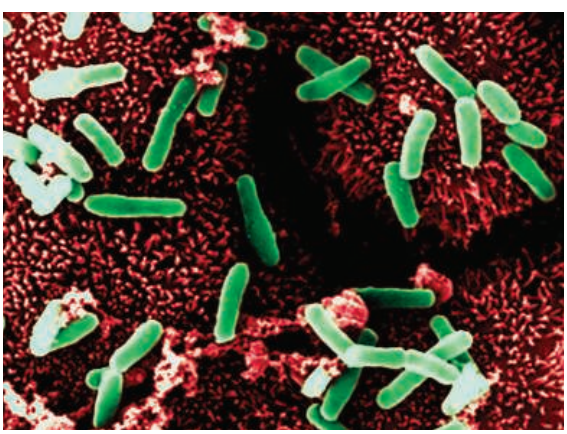
With reporting by Ann Gibbons.

INFECTIOUS DISEASE

RU-486–Linked Deaths Open Debate About Risky Bacteria

ATLANTA, GEORGIA—Government officials and scientists convened last week to address troubling questions about two deadly types of bacterial infections that may be growing more common. One pathogen, *Clostridium sordellii*, has drawn intense political and scientific interest after being linked to deaths in young women following medical abortions, most with the abortion pill RU-486. The other, its cousin *Clostridium difficile*, is a growing scourge in hospitals.

The meeting, held at the Centers for Disease Control and Prevention (CDC) here, was called three months ago largely because of the abortion-associated deaths, which then stood at five and are now thought to number seven. But it turned into a much broader, handwringing discussion over how much remains to be learned about both types of *Clostridium*.



Gutted. *Clostridium difficile* bacteria (green-white), attached here to human intestinal tissue, are making more and more people sick.

C. difficile, which ravages the colon, has killed hundreds of hospital patients since 2000 and is increasingly showing up in healthy people and in animals. That's led to some concern about transmission through the food chain. Indeed, seven *C. difficile* patients appear to harbor animal strains of the bacterium, Clifford McDonald, a medical epidemiologist at CDC, announced. "We are a little disturbed" by that, he says.

Rates of *C. difficile* infections have soared recently, doubling in U.S. hospitals between 2000 and 2003 and jumping another 25% in 2004. In the United Kingdom, the disease rate leapt from 1 in 100,000 people to 22 in 100,000 over 10 years. The mortality rate also appears to be increasing, from about 1% to almost 7% in some cases, such as an epidemic in Quebec hospitals in Canada 2 years ago.

Typically, *C. difficile* sickens hospital patients who have taken antibiotics, although how the drugs predispose patients to the germ is "pretty much a black box," says Ciarán Kelly, a gastroenterologist at Harvard's Beth Israel Deaconess Medical Center in Boston. The bacterium is showing up more and more outside hospitals and among people with no recent antibiotic exposure. Analyzing the responsible strains, says McDonald, could sort out whether *C. difficile* has become food-borne. It could also determine whether the bacterium has mutated. In December, scientists described in the *New England Journal of Medicine* a novel strain of *C. difficile* that may churn out more toxin.

A better grasp of the bacterium's basic biology could offer clues to preventing and treating infections, but working with the microbe is a challenge. *C. difficile*, like *C. sordellii*, is difficult to manipulate genetically. And some work suggests that the bacterium behaves differently in humans than in animals, implying that animal models may be misleading, says Kelly.

C. sordellii is even less well understood. It's not clear what predisposes people to an infection, and unlike *C. difficile*, which may respond to antibiotics, *C. sordellii* infections are rarely treatable. When the U.S. Food and Drug Administration (FDA) approved RU-486 in 2000, it urged that possible adverse events be reported to the agency, and it was those reports, of a handful of women succumbing within hours or days to a terrifying infection, that first alerted health officials. The overall risk of death from these infections has been estimated at about 1 in 100,000.

Some have speculated that vaginal rather than oral administration of misoprostol—a drug that acts with RU-486 to induce an abortion—was a factor in the deaths. But the meeting underscored that "this is a far more complex medical and epidemiologic situation than originally appeared to be the case," says Sandra Kweder, deputy director of FDA's Office of New Drugs.

CDC's Marc Fischer detailed 10 fatal cases of *C. sordellii* genital tract infection from 1977 to 2001, which the agency found by combing through old records. Eight cases occurred after women gave birth; one followed a miscarriage; and the last was not associated with pregnancy. McDonald presented four additional cases CDC is investigating, three of which are thought to have followed nonsurgical abortions and a fourth following a miscarriage. To make matters more confusing, two of those new cases involve not *C. sordellii* but a third member of the clostridium family, *C. perfringens*. "If you look at the presentation of these illnesses, they always come after delivery, after miscarriage, after the passage of abortion," says David Soper, an obstetrician-gynecologist at the Medical University of South Carolina in Charleston. "Does pregnancy hold the key?"

Esther Sternberg of the National Institute of Mental Health in Bethesda, Maryland, has found that *C. sordellii* toxins disrupt hormone receptors for glucocorticoids, which may predispose women to an excessive inflammatory response in the presence of the bacteria. Fischer noted that CDC is limited in its ability to track down old *C. sordellii* cases. It is asking physicians to report suspicious deaths following pregnancy or miscarriage.

Companies are now developing vaccines, as well as drugs that bind to *C. difficile*'s toxins. A National Institutes of Health (NIH) official at the meeting urged attendees to submit *Clostridium* research proposals. Meanwhile, CDC, FDA, and NIH plan to identify research priorities in the field. So far, FDA has given no indication that it will change how RU-486 is marketed.

–JENNIFER COUZIN

CREDIT: MED. MIC. SCIENCES CARDIFF UNI/WELLCOME PHOTO LIBRARY

SCIENTIFIC MISCONDUCT

Invention of China's Homegrown DSP Chip Dismissed as a Hoax

In a major embarrassment for China's national electronics R&D program, an inventor's claim to have created a series of homegrown computer chips has been declared a fraud. After a months-long investigation, Shanghai Jiao Tong University (SJTU) announced on 12 May that it found "serious falsification and deception in the research and development of the Hanxin series of chips led by [SJTU dean] Chen Jin." The university announced that Chen had been dismissed. Chen did not respond to telephone or e-mail messages.

Chen won national acclaim in February 2003 when he unveiled what he described as the first digital signal processor (DSP) chip designed and manufactured in China, called Hanxin-1 or "Chinese chip." He quickly followed with two improved designs and promised a fourth and fifth generation with both a DSP and a central processing unit. The 37-year-old inventor built his career on aiming, as he told a reporter, "to put the label 'Made in China' on high-end computer chips."

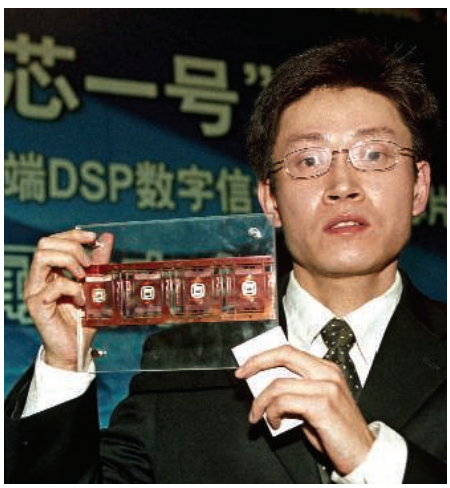
With a 1998 Ph.D. in computer engineering from the University of Texas, Austin, Chen spent a short stint as a test engineer at Motorola Semiconductor Product Sector in Texas, now called Freescale Semiconductor, and returned to China in 2000. At SJTU, Chen embarked on a road to take back China's DSP market shares. In less than 2 years, he managed to set up an integrated-circuit design lab and had a product ready.

Government and academic leaders embraced the inventions. Chen was appointed dean of SJTU's newly established School of Microelectronics; he founded the company SJTU HISYS Technology Ltd. and became its CEO. More than $7 million in public R&D funds poured in. The Shanghai government named Chen CEO of Shanghai Silicon Intellectual Property Exchange, a platform established in 2003 with $3.75 million in municipal funds for trading semiconductor rights.

But on 17 January, an anonymous posting on a Chinese Web site presented evidence alleging that the project was a fraud. The tipster claimed that Chen had purchased 10 Motorola DSP chips in August 2002 and had the original logo sanded off and replaced with HISYS and SJTU labels. According to the allegations, Chen promoted the chips as his Hanxin-1 design and later passed off other derivative products as his own inventions.

HISYS Technology issued a statement on 21 January calling the allegations "pure fabrication." However, 5 days later, SJTU issued a statement expressing concern over the alleged fraud and announcing that the university had asked national ministries and the Shanghai government to help investigate.

The investigation was organized by China's Ministry of Science and Technology (MOST)—a major investor in the project—the Ministry of Education, and the Shanghai government. An expert team interviewed Chen, the still-anonymous Internet tipster or tipsters, and others. It inspected and compared technical documents on site and checked the design and process specifications of Hanxin chips 1 through 4.



Before the fall. Microelectronics wizard Chen Jin with his chip at a press conference in February 2003.

CREDIT: AP PHOTO

Last week, SJTU released the team's findings: The device Chen had displayed as Hanxin-1 at a press conference in 2003 was not the one that had been submitted for evaluation; instead, Chen substituted another chip that his lab did not design. SJTU's report also said that Chen did not own the "core technology" of other chips that he claimed. Chen, the report said, "used false results to cheat evaluation experts, Shanghai Jiao Tong University, his research team, local government, ministries of the central government, as well as the media and the public," but the report does not say how the evaluation experts were cheated. MOST terminated Chen's ministry-funded projects and asked him to return the research funds.

Chen appears to be moving on to other ventures. At a low-key news conference last month, he announced that HISYS Technology—now severed from the university—is forming an alliance with Skyworks Shanghai to develop products for the mobile phone market.

–HAO XIN

It's Unanimous: Patents Not Omnipotent

Patent holders don't have an automatic right to shut down their competitors to protect their intellectual property rights, the Supreme Court decided this week. The 15 May ruling involves a case brought by a small Virginia company, MercExchange, against eBay, the online trading giant. The high court focused on the proper use of injunctions, the orders that judges file to halt companies from operating after they're found guilty of infringement.

In 2003, after eBay was found guilty of infringement, a district court denied MercExchange's request to halt the operation of aspects of its rival's Web site, citing fairness. But last year an appellate court declared that judges should deny injunctions only in "exceptional circumstances." In sending the case back to the lower court, however, the high court slammed the appeals court for having "erred in its categorical grant" of injunctions for patentees. Yet the nine justices also recognized that "university researchers or self-made inventors" can stop infringers' operations even if as innovators they don't market their technology.

"If this were golf, this [ruling] was right down the fairway," says Kevin Noonan, a patent attorney with McDonnell Boehnen Hulbert & Berghoff LLP in Chicago, Illinois. Information technology companies and the biotech and pharma communities line up on opposite sides of the issue, with the former complaining about so-called "patent trolls" who target their products and the latter worried about losing their research investments. Legislation to reform the patent system is stalled in Congress.

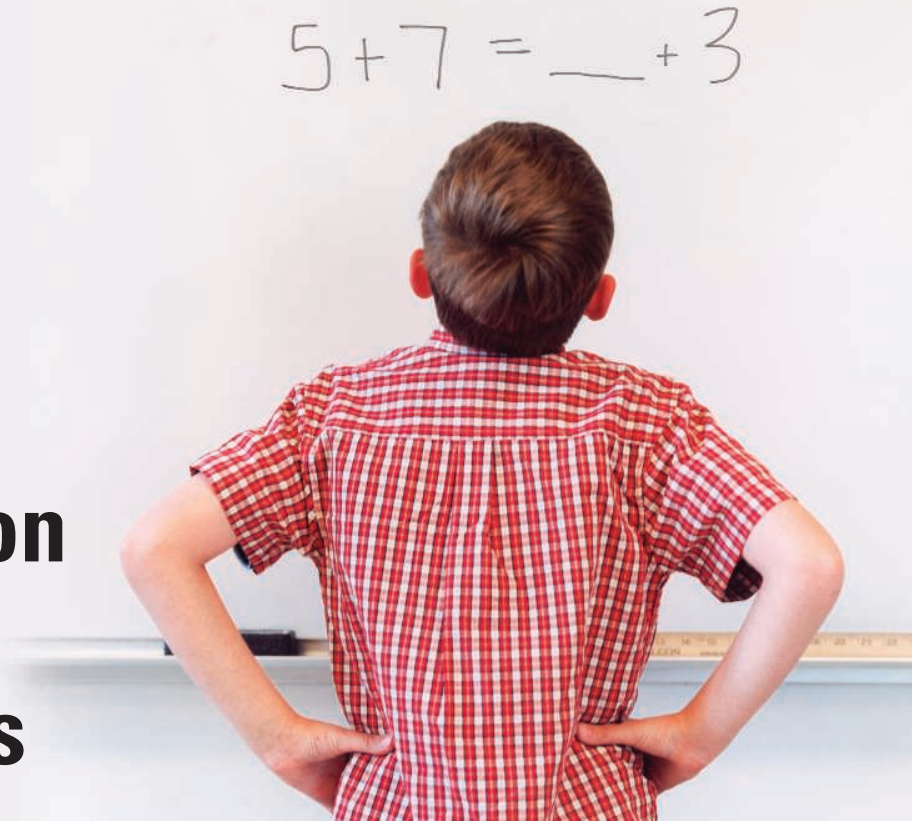
–ELI KINTISCH

Government Crackdown, Please

Concerned about the chaotic way scientific misconduct allegations in China are being publicized, Chinese scientists are asking the government to step in. More than 120 Chinese researchers, most U.S.-based, have signed a letter calling on research agencies to create an official process for addressing such charges. The letter was drafted by Xin-Yuan Fu, a microbiologist at Indiana University School of Medicine in Indianapolis, who says the "rule of law" should also apply to Chinese science. The appeal was triggered by a heated Internet debate on the credibility of two biomedicine papers by Wei Yuquan, vice president of Sichuan University in Chengdu. Wei, 46, who specializes in tumor biotherapy, has denied wrongdoing and called for an investigative hearing.

–HAO XIN

Published by AAAS



For years, mathematicians and math educators have blamed one another for the inadequacies of U.S. mathematics education. But both sides may finally be headed toward agreement on how to fix the system

Finding Common Ground in the U.S. Math Wars

LIKE A SHERIFF SUMMONED TO RESTORE order to a lawless town in the Wild West, Richard Schaar knew that taking on the Math Wars would be a rough assignment. An applied mathematician and former president of the calculator division at Texas Instruments (TI), Schaar was part of an industry-led panel trying to improve U.S. science and math education a few years back when he realized that a huge schism in the community would likely block any effort to reform elementary and secondary school mathematics.

"I hate labels, but in general the professional mathematicians were on one side, and the math educators were on the other," says Schaar, describing a debate, triggered by a huge backlash to a 1990s reform movement, that has persisted despite mounting concern about how poorly U.S. students fare in international comparisons. "The argument over direct instruction versus discovery learning, as the two sides are commonly described, was pulling the field apart. The mutual respect had gone away. And in that climate, any attempt to improve math standards at the state level would have been doomed to failure."

The solution seemed obvious to him: Bring together a handful of top guns from each side and hope for harmony rather than bloodshed. And that's exactly what Schaar has done, in the Common Ground initiative (www.maa.org/common-ground). The six-member group has made modest but impressive progress over the past 18 months in finding agreement on issues that for the last decade have led mathematicians and math educators, in the words of one mathematics society executive, "to sit on the sidelines and lob bombs at each other." (To be fair, both sides claim to be appalled by the analogy to warfare. But they use combat imagery repeatedly in conversations as a shorthand to describe their experiences.)

The Common Ground initiative is one of several hopeful signs that the two sides may be ready to call a truce and work together to improve U.S. mathematics education. Last month, the country's largest group of mathematics educators, the National Council of Teachers of Mathematics (NCTM), endorsed a short list of math skills, by grade, that every elementary and middle school student needs to master. These skills, called Curriculum Focal Points, are an attempt to correct what math educators decry as "mile-wide, inch-deep" curricula in most U.S. schools that leave many students unprepared for high school and, ultimately, precludes them from pursuing careers in science and engineering. This week, the Department of Education named mathematicians, educators, and community leaders to a presidential panel that will review the state of mathematics education (see p. 982). Observers are hopeful that the easing of tensions will improve the quality of the panel's recommendations on bread-and-butter issues such as student instruction, teacher training, and the additional research needed to enhance each area, not to mention make those recommendations easier to sell.

"I think Common Ground is a historic and groundbreaking exercise," says Frances "Skip" Fennell, a mathematics education professor at McDaniel College in Westminster, Maryland, and NCTM president. "I worked in the education directorate at NSF [National Science Foundation] in the late 1990s, and I was blown away by the anger in the community. This is exactly what we need to get things moving forward."

All for algorithms

Professional mathematicians blame themselves for some of those angry words. They were heavily involved in a major reform of the U.S. mathematics curriculum in the 1960s, after Sputnik, that was widely criticized as too difficult for the average student. In response, mathematicians largely withdrew from the fray and were silent when math educators promulgated the next round of reforms in response to a 1983 report that said low student achievement in reading and math was putting the country at risk. "There's been a divide between education and subject matter fields for a long time, but it's had its worst consequences in math," notes Roger Howe, a Yale University mathematician who has thought hard about the mathematical foundations of elementary principles such as place value. And when the mathematicians belatedly discovered aspects of the new courses that they didn't like, they unleashed their wrath upon federal officials and math educators, castigating them at every opportunity for demanding too little of students and watering down their discipline.

Given the rancorous tone of the debate, Schaar knew that he needed to sign up leading figures from both sides. He spent a year picking his team: two mathematics professors who have been sharp, public critics of the reform curricula (R. James Milgram of Stanford University in Palo Alto, California, and Harvard University's Wilfried Schmid) and three math educators in the forefront of those reforms

CREDIT: ADAPTED FROM LISA PINES/GETTY IMAGES

Published by AAAS

(Deborah Loewenberg Ball of the University of Michigan, Ann Arbor; Joan Ferrini-Mundi of Michigan State University in East Lansing; and Jeremy Kilpatrick of the University of Georgia, Athens). In December 2004, the same month he retired from TI, Schaar convened the first meeting of the Common Ground initiative, with himself as facilitator.

Six months and six meetings later, the group issued a three-page document describing a handful of principles that should guide math education from kindergarten through high school. The principles include the automatic recall of basic facts, the importance of abstract reasoning, the need to acquire a mastery of key algorithms, and the judicious use of calculators and real-world problems. Two months ago, an expanded group met for a weekend to tackle the topics in greater detail, and last week, initial working papers from that meeting were posted. The core group met again last weekend to plot its next steps, as well as to clarify its earlier statement about setting high expectations for students—one that's been misinterpreted as an argument for making calculus a required course in high school.

The document doesn't say when or how any of the concepts should be taught. Common Ground is not a curriculum, Schaar points out. The most its participants can hope to achieve is to influence the process by which states develop standards, adopt textbooks, and develop the assessment tools to measure what students should be learning. Even so, their carefully worded statements on selected topics reflect hard-fought compromises on core issues that have roiled the community for more than a decade and that, once resolved, could pave the way for continued progress.

"There will always be differences," says Milgram, who in 2000 testified before Congress that "the sad state of U.S. mathematics education" is the result of "a constructivist philosophy" promoted by NCTM standards and endorsed by NSF and the Department of Education, the two leading federal sources of support for teaching mathematics. "But if we can agree on the essential content that students need to know, then the other fights become manageable. And I'd say that there has been far more agreement than disagreement."

Ball, who has done pioneering work on what math teachers need to know to do their jobs well (i.e., not just how to teach long division but also to understand why Susie's method is incorrect), believes that the process has been just as important as the product. "Our goal was to provide leadership to the field, to say to everybody: 'If we can do it, then the rest of you can, too.' And I think we've shown that it's possible to come together on many of the flash points."

One major flash point is the use of algorithms—how to do long division, for example—and the memorization of the facts upon which they are based. Many mathematicians maintain that current state standards and instructional materials downplay the use of such time-tested algorithms or allow students to bypass them entirely by using calculators. So when Common Ground asserts that "students should be able to use the basic algorithms of whole number arithmetic fluently, and they should understand how and why the algorithms work," the participants are trying to stitch up a vast rift in the community.

"Of course kids have to know how to compute and know their basic facts. But they also have to make sense of what they are being taught and explore the ideas with open-ended problems," says Sybilla Beckmann Kazez, a mathematician at the University of Georgia, Athens, who is well respected by both camps. "If you put it that way, everybody would agree." Schaar concurs that the initiative has only scratched the surface on this contentious subject: The question of algorithms "is an incredibly challenging area that will require additional exploration."

Getting to the (focal) point

NCTM's new curriculum focal points, covering prekindergarten through grade eight, are also just beginning their long journey through the educational system. (The document won't even be released publicly until fall, officials say, although drafts have circulated and the council's executive board approved the latest version last month at the organization's annual meeting in St. Louis, Missouri.) With three per grade, the focal points address what math educators decry as overly broad and shallow curricula in most U.S. schools that hinder mastery and prepare students poorly for college-level work.

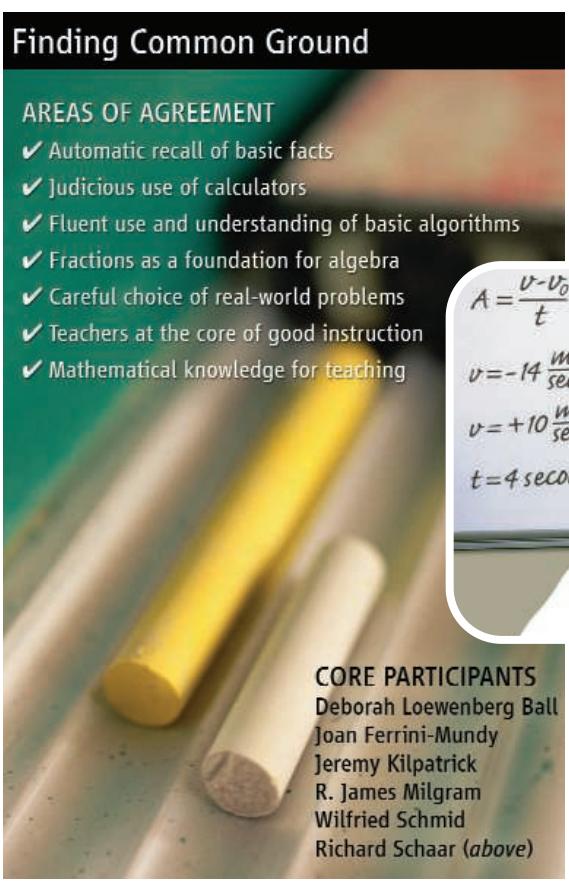
NCTM President Fennell says the focal points are intended to provide "curricular relief" to elementary and middle school teachers whose school districts expect them to achieve as many as 100 objectives in mathematics. Many of those objectives span several grades, with teachers expected to tailor them to the maturing child. But there's no urgency because teachers know that their students will get another bite of the apple the following year.

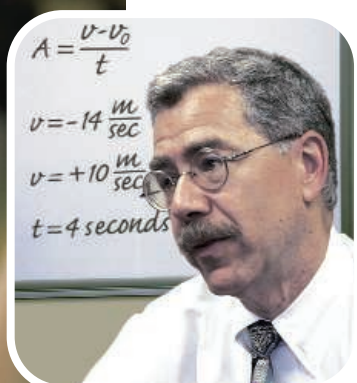
"While lots of things are important, we're saying to teachers that here are three things you need to zero in on," says Fennell. "For example, we'll teach some probability in the fourth grade. But it's not as important as multiplication," which takes center stage alongside fractions and decimals and the concept of area. Second graders should concentrate on addition and subtraction, place value, and linear measurement, says NCTM, even if their teachers also touch upon other topics.

Although focal points must first be woven into state and district guidelines to have any real effect, the council's action already represents a significant move toward common ground: Professional mathematicians love to attack the 1989 and 2000 NCTM standards, and they see focal points as a tacit admission that some of their criticisms were on the mark. They also welcome the message that, for most students, less is more.

"The idea of coming up with a few topics that should be addressed in K through 8 is a very needed step," says Richard Askey, a professor emeritus of mathematics at the University of Wisconsin, Madison, and an outspoken critic of earlier NCTM standards and curricula based on them. "I think that publishers, who now have to deal with all [different] state standards, will also like the idea" of a limited number of key objectives for each grade.

Jane Schielack, a mathematician and math educator at Texas A&M University in College Station who led the NCTM task force that assembled the focal points, agrees that they are very much a product of the





CREDITS: GETTY IMAGES; (INSET) COURTESY TEXAS INSTRUMENTS INC.

Published by AAAS

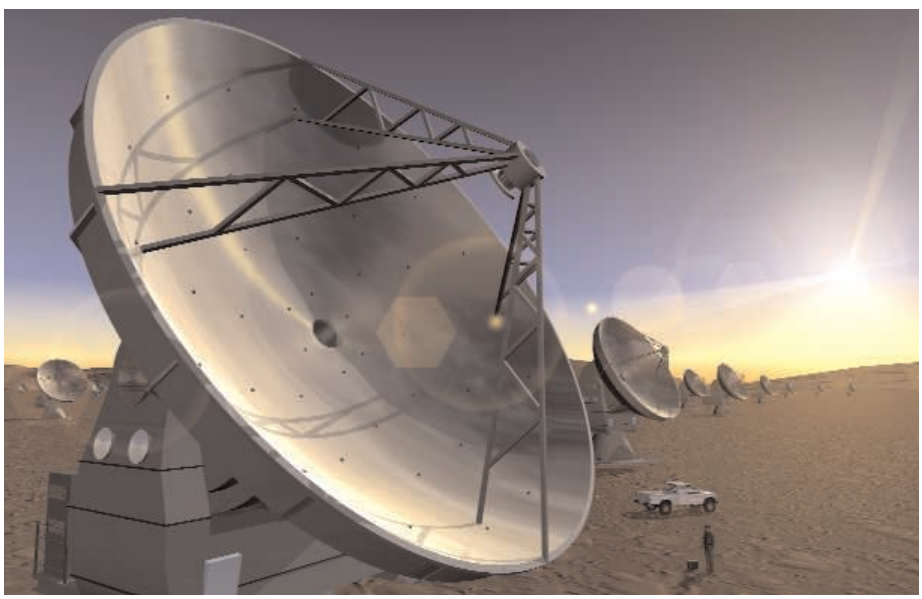
times. "This is something we couldn't have done 4 or 5 years ago," she says. In addition to the greater emphasis on accountability spawned by the 2001 federal No Child Left Behind law, Schielack cites the growing recognition that some countries, notably Singapore and China, excel on international student comparisons because of a national curriculum that focuses on a small number of topics and policies that give teachers the necessary training and resources to get the job done. "That's the biggest difference between the United States and the top-achieving nations," agrees Milgram. "Having NCTM come out with a statement to this effect should make an enormous difference on what we expect kids to learn."

Even so, nobody expects Common Ground and focal points, by themselves, to usher in a golden age of quality mathematics education. There's too much that remains to be done. "It's a long, long journey," says Hung-Hsi Wu, a mathematician at the University of California, Berkeley, who runs summer institutes for classroom teachers whose grasp of basic mathematics is often poor or nonexistent. "Better mathematics education in the United States won't take place in the next 10 years. I think it will take 30 years."

At the age of 60, Schaar doesn't plan on staying in the line of fire for quite that long. But he's not ready to saddle up and ride out of Dodge. Schaar believes that Common Ground, funded by NSF and TI and staffed by the Mathematical Association of America, has restored a measure of civility to the debate. And this month, after a coalition of 16 leading mathematical societies applauded his 2-hour presentation and told him to keep up the good work, he said that kind of support is exactly what's needed.

"I'm not looking for an endorsement," he says. "I'm looking for help in getting more people involved." A bigger "in" crowd means fewer outcasts. And that's good news for a sheriff.

–JEFFREY MERVIS



All together now. For different observing jobs, ALMA's 50 antennas can be rearranged with a giant purpose-built truck.

CREDIT: H. ZODET/ESO

ASTRONOMY

After a Tough Year, ALMA's Star Begins to Rise at Last

Cost hikes, scarce labor, and management changes have buffeted the first global telescope array, but new funding agreements may augur smoother sailing ahead

The world's largest ground-based astronomy project, the Atacama Large Millimeter Array (ALMA), is back on track after a tumultuous couple of years that have seen costs balloon by about 40% and the capability of the enormous microwave telescope scaled back.

ALMA, with an overall budget now in the region of $1 billion, is a collaboration between the United States, the European Southern Observatory (ESO), and Japan, plus minor partners Canada and Spain. As a result of skyrocketing prices in commodities needed to build its antennas and huge hikes in labor costs in Chile, where ALMA is being built, astronomers have had to go cap in hand to their funders for more money. ESO agreed to swallow its share of the increases last autumn, but it was not until last week that the U.S. National Science Foundation (NSF) won agreement from its governing board. "It's been a fairly intense 18 months," says astronomer Christine Wilson of McMaster University in Hamilton, Canada, chair of ALMA's scientific advisory committee.

"I'm told that most big projects go through something like this," Wilson says. "Cost increases are a given." But for researchers waiting to see whether funders would keep faith with the project, the process has been nerve-wracking. "We were holding our breath back in the summer and fall for ESO," Wilson says. "It's been a very stressful situation for everyone in the project." U.S. team members had to await the outcome of a series of cost reviews, but in a meeting on 10 May, the National Science Board gave NSF permission to increase U.S. spending on ALMA from $344 million to $499 million, subject to the approval of Congress. According to ESO's Thomas Wilson, European project scientist on ALMA, during these discussions there was an unspoken warning from the funders: "This is it. Don't come back and ask for more."

ALMA, the first truly global effort in ground-based astronomy, grew out of three separate projects. U.S. astronomers started discussing a Millimeter Array in the mid-1980s; European plans for a Large Southern Array took shape about a decade later. ESO and the U.S. National Radio Astronomy Observatory (NRAO) in Socorro, New Mexico, began discussions on merging the two projects in 1997 and in June 1999 agreed to build a joint instrument comprising 64 12-meter antennas spread over an area up to 12 kilometers across. The array took its new name from Chile's Atacama desert, where researchers had found a wide plateau, the Llano de Chajnantor, which at 5000 meters altitude is high enough and dry enough to avoid most of the atmospheric water vapor that blocks signals at the wavelengths ALMA is designed to receive.

The push for such an instrument came because better receivers, fast digital electronics, and antenna design were improving the capabilities of millimeter-wave telescopes. Astronomers calculated that a large number of receivers arranged as an interferometer could rival the resolutions of the best optical instruments, such as Hubble and ESO's Very Large Telescope in Chile. At millimeter and submillimeter wave-

lengths, astronomers can study the lowest-energy emissions from simple molecules. With ALMA, they hope to peer into star-forming galaxies when the universe was young to see whether stars formed in a burst early on or more steadily over a long period. Closer to home, they can see whether disks of dust and gas around young stars—places where planets could form—are commonplace or rare.

Japan, which had been developing its own Large Millimeter and Submillimeter Array, joined the club in 2001. The plan is for Japan to construct a parallel instrument, the Atacama Compact Array (ACA), made up of 4 12-meter antennas and 12 7-meter antennas. Sited next to the main array, ACA will be better able to image extended diffuse objects. In addition, Japan is providing receivers to cover three extra wavebands for antennas in both ACA and the main array.

At first, everything moved along according to plan. Prototype antennas for the main array were ordered from two suppliers, one in Europe and one in the United States. Work crews began preparing the site at Llano de Chajnantor in late 2003. Once delivered, the prototype antennas were put through a series of tests at a specially built facility in Socorro, home of the Very Large Array radio telescope. Testing was completed in April 2004 with a view to awarding the antenna contracts—the biggest items on the ALMA shopping list—later that year.

ALMA researchers, however, were not happy. "The first round of tests were not conclusive," says Thijs De Graauw of SRON, the Netherlands Institute for Space Research, and chair of ALMA's management advisory committee. "There were valid concerns," adds astronomer Lee Mundy of the University of Maryland, College Park. "They were asking for a very precise antenna and wanted to make sure it could accomplish the science."

New tests were ordered, but the delay proved costly. At the time, the prices of commodities essential for the antennas' construction, such as steel, were going through the roof. And as the extra tests dragged on into 2005, ALMA managers had to ask the manufacturers to resubmit their bids for building the production antennas. The bids came in much higher than managers had expected and threw the project into crisis. Asked whether ALMA could make do with fewer antennas, the scientific advisory committee concluded that the array could achieve its primary science goals with 50 rather than 64 dishes, but observations would take longer and would be more prone to systematic errors. An array of less than 50 instruments would still be "a superb instrument," the advisers said, but its goals would be compromised.

"We decided to reduce the number of antennas so the cost increase would not be too large," says ESO Director General Catherine Cesarsky. The North American team went ahead in July 2005 and placed an order for 25 antennas, with an option to buy another seven. ESO was poised to follow suit, but then it hit another snag. Under its rules, it had to take the lowest bid that met specifications. ESO had planned to buy from the same company NRAO had ordered from, VertexRSI of Kilgore, Texas. But the European consortium led by French-Italian company Alcatel Alenia Space submitted a cheaper revised bid. Before signing on the dotted line, Cesarsky says ESO waited to see a cost review of the whole ALMA project that was completed in October and carried out a review of all its programs to see whether enough economies could be made to cover the extra costs.

Concerns remained even after ESO ordered its 25 antennas from Alcatel last December. Some researchers worried that having two sets of antennas from different suppliers would increase costs down the line because it would require double the number of technicians and spare parts. But in January, a "delta" review of the increased cost reported that it was unlikely to be more than 1% of ALMA's total budget. Meanwhile, other costs were also draining ALMA's coffers. Chile's economy has been booming, and the consequent boost to the construction industry has made labor hard to find and more expensive. In addition, copper prices are at an all-time high, and northern Chile has extensive copper deposits. Chilean workers, it turned out, would rather mine copper than work in the cold airlessness of 5000 meters.

Labor troubles have exacerbated another hurdle ALMA is working to overcome: learning to manage a global engineering project. "Astronomers are not used to this scale of project," Mundy says. "It's taking astronomy into the big league." Some have charged that managers' cost estimates at the start of the project were unrealistic and that ESO based its estimated construction costs on the other observatories it had built in Chile, which were all at lower altitudes. "Assumptions were optimistic," says De Graauw. "Errors came from not knowing in enough detail what was to be built." Says Mundy: "In a project of this scale, managers and management systems are needed. These were not components of the original pricing."

Cesarsky acknowledges that running the project with two management teams separated by the Atlantic has been difficult: "It was not clear who should make decisions. A strong central management was needed." More control has now been put in the hands of the Joint ALMA Office in Santiago, Chile's capital, Cesarsky says.

The flurry of reviews that have assessed the project from within and from outside have now given it a clean bill of health. "I think things are going along very well," says Al Wootten, ALMA's North America project scientist. But for researchers, the necessity to cut back the number of antennas to 50 rankles. "People are unhappy about it still," says ESO's Wilson. Cesarsky thinks there's still a possibility that the array can be built at full strength, "if we're lucky and have not spent our contingency." Not everyone is so positive. "Do we skimp and endanger the whole instrument? Surely it's better to do it right once," argues Mundy. "I haven't heard any way to get there, but the door is still open."

–DANIEL CLERY

CREDIT: HANS HERMANN HEYER/ESO

High and dry. With a Chilean construction boom in progress, ALMA managers are having trouble finding people to work in the thin desert air 5000 meters above sea level.

Published by AAAS

Speed matters. It has taken just over 5 years and $37 million to complete China's new tokamak, according to the Institute of Plasma Physics.

ENERGY ALTERNATIVES

Waiting for ITER, Fusion Jocks Look EAST

China is breaking new ground with a fusion test bed that will tide researchers over until the ITER megaproject comes online

HEFEI, CHINA—The official launch of the International Thermonuclear Experimental Reactor (ITER) project next week will mark a coming of age for fusion research in Asia. When the $11 billion effort was initiated in 1985, ITER's four original backers—the United States, the European Union, Japan, and the Soviet Union—accounted for nearly all worldwide research into harnessing fusion, the process that powers the sun, to produce energy. But now the three newest ITER partners, China, South Korea, and India, are showing that they didn't just buy their way into one of the biggest physics experiments since the Manhattan Project: They are contributing crucial expertise as well.

The first new Asian fusion tiger out of the gate is the Institute of Plasma Physics (IPP) of the Chinese Academy of Sciences, which in March completed testing a machine that has never been built before: a fully superconducting tokamak. This toroidal vessel isn't the largest or most powerful device for containing the superhot plasma in which hydrogen isotopes fuse and release energy. But until India and South Korea bring similar machines online (see sidebar, p. 993), it will be the only tokamak capable of confining a plasma for up to 1000 seconds, instead of the tens of seconds that machines elsewhere can muster. ITER, expected to be completed in Cadarache, France, in 2016, will have to sustain plasmas far longer to demonstrate fusion as a viable energy source. But researchers from China and around the world will be able to use IPP's Experimental Advanced Superconducting Tokamak (EAST) to get a head start on learning to tame plasmas for extended periods. "This will make a big contribution for the future of fusion reactors," declares Wan Yuanxi, a plasma physicist who heads EAST.

Fire when ready. EAST will fill a crucial gap for fusion researchers until ITER is built, says Director Wan Yuanxi.

Fusion research over the next decade will be probing the physics of steady-state plasmas like those promised by ITER, says Ronald Stambaugh, vice president for the Magnetic Fusion Energy Program at General Atomics in San Diego, California. "EAST will play a big role in that," he says. Others credit IPP for building its advanced tokamak fast, in just over 5 years, on a shoestring $37 million budget. That's a fraction of what it would have cost in the United States, says Kenneth Gentle, a plasma physicist and director of the Fusion Research Center at the University of Texas, Austin. "That they did this in spite of the financial constraints is an enormous testimony to their will and creativity," adds Richard Hawryluk, deputy director of the Princeton Plasma Physics Laboratory.

IPP adroitly fills a generational gap. Fusion power will rely on heating hydrogen isotopes to more than 100 million degrees Celsius, until they fuse into heavier nuclei. The leading design for containing this fireball is the tokamak, a doughnut-shaped vacuum chamber in which a spiraling magnetic field confines the plasma. Ringlike metal coils spaced around the doughnut—toroidal field coils—and a current in the plasma produce this spiraling field. Additional coils in the center of the doughnut and along its circumference—poloidal field coils—induce the current in the plasma and control its shape and position.

Early tokamaks had circular cross sections and copper coils, which can only operate at peak power in brief pulses before overheating. ITER will be far more sophisticated. It will have a D-shaped cross section, designed to create a denser plasma that can generate its own current to supplement the induced current, reducing energy input. And coils will be superconducting. (No major tokamak has had superconducting poloidal field coils.) At temperatures approaching absolute zero, superconductors carry current without generating resistance, allowing more powerful magnetic fields that can be maintained longer.

Researchers want to try out a D-shaped, fully superconducting test bed before scaling up to ITER, which will be two to three times the size of current tokamaks. The Princeton Plasma Physics Laboratory had planned to build such a device. But a cost-conscious U.S. Congress killed their $750 million Tokamak Physics Experiment in 1995. EAST and the two other Asian tokamaks under construction intend to fill this gap.

"We recognized this was an opportunity for us to make a contribution for fusion research," Wan says. For support, he tapped into China's worries about its growing demand for energy. "There is no way we can rely entirely on fossil fuels," he says. China's government approved EAST in 1998.

IPP faced an enormous challenge. The institute, founded in 1978, had built a few tiny tokamaks in the 1980s and got a hand-me-down, partially superconducting tokamak from Russia's Kurchatov Institute in 1991. EAST would be a totally different beast. "We didn't have any experience in the design, fabrication, or assembly of these kinds of magnets," Wan admits. Neither did Chinese manufacturers.

Industrial partners supplied parts of the tokamak, including the vacuum vessel. But the superconducting coils and many other high-tech components would have been too expensive to import. "We had to do [these] ourselves," says the tokamak's chief engineer, Wu Songtao. So Wu's team bought precision milling machines, fabricated their own coil winders, and built a facility to test materials and components at cryogenic temperatures. "They literally built a whole manufacturing facility on site," says Hawryluk.

CREDITS (TOP TO BOTTOM): DONG YIDONG; D. NORMILE/*SCIENCE*

Published by AAAS

IPP physicists and engineers passed a major milestone earlier this year, when they tested the entire assembled device, cooling the 200 tons of coils to the operating temperature, 4.5 kelvin. They discovered only minor, fixable glitches, Wan says, and are now undertaking the necessary tweaks and installing shielding materials and diagnostic devices. In August, they plan to inject hydrogen and fire up EAST's first plasma.

With the tokamak passing its cool-down test, Wan says the team was "finally able to get a good night's sleep." They are now planning experiments to explore how to control D-shaped plasmas. Tugging a plasma into a specific shape can create instabilities, Gentle says. Control is all the more difficult because superconducting coils respond poorly to current fluctuations. IPP will probe these issues. "That's where the science is going to be extremely valuable," says Hawryluk.

EAST has limitations. The most significant is that, unlike ITER, it will not attempt a burning plasma, in which at least half the energy needed to drive the fusion reaction is generated internally. ITER will use a combination of deuterium and tritium (hydrogen isotopes with, respectively, one and two neutrons in the nucleus), which fuse at a lower temperature than other gases, to achieve a burn. Because radioactive tritium requires specialized and expensive handling systems and shielding, EAST will use only hydrogen or deuterium.

That limitation is hardly dampening enthusiasm for the hot new kids on the block. IPP researchers, says Hawryluk, "have already put themselves on the fusion community map."

–DENNIS NORMILE

With reporting by Gong Yidong.

Asian Fusion

India, Korea, and possibly Japan are joining China in building next-generation tokamaks. These machines seek to fill a research gap on the road to the International Thermonuclear Experimental Reactor (ITER) by employing all-superconducting coils to study the physics of confining plasmas for long durations, which current tokamaks can't do.

■ India's Institute for Plasma Research is now commissioning its Steady State Superconducting Tokamak. An engineering test at cryogenic temperatures turned up problems that are now being addressed. Institute plasma physicist Y. C. Saxena says they are hoping to try a second engineering test later this month. If that goes well, they will attempt their first plasma in the summer. The $45 million project, launched in 1994, is the smallest of the new tokamaks. But Saxena says they believe they can help unravel the physics of long-lasting plasmas.

■ The most ambitious machine is the Korean Superconducting Tokamak Reactor (KSTAR), being built by the National Fusion Research Center in Daejeon. KSTAR relies on superconductors made from the more advanced niobium-tin alloy that ITER will employ. The $330 million project was delayed because of Korea's late-1990s economic crisis. Project Director Lee Gyung-su says they are now aiming for first plasma in early 2008.

■ For several years, Japan's Atomic Energy Agency has been studying the possibility of upgrading its JT-60 tokamak to be fully superconducting. Japan may get funding for the upgrade from the European Union as compensation for its assent on the agreement to build ITER in France. An agency spokesperson says key decisions are under negotiation. **–D.N.**

SCIENTIFIC OPENNESS

Should Academics Self-Censor Their Findings on Terrorism?

Some government-funded researchers believe their papers require special handling. But others say that creating such a gray area undermines academic freedom

Last year, after Detlof von Winterfeldt and his colleagues at the University of Southern California (USC) in Los Angeles finished a study on the likelihood and impact of a dirty bomb attack by terrorists on the Los Angeles harbor, they omitted some important details from a paper they posted on the Internet. Although the team had used no classified material, von Winterfeldt felt that self-censorship was prudent given the subject matter. It's also in line with draft guidelines being considered by the U.S. Department of Homeland Security (DHS), which funds the Center for Risk and Economic Analysis of Terrorism Events that he directs. "We were still able to present the methodology behind the analysis fully and effectively," he says. "It made perfect sense to make those changes."

But some scientists say that stance conflicts with academic freedom, and that the public deserves access to anything not explicitly classified. They worry that the actions of the USC researchers could serve as a model for restricting the conduct and dissemination of university research. Their concerns are tied to an ongoing effort by the Bush Administration to draw up common standards across federal agencies for withholding information under the rubric of sensitive but unclassified (SBU) material.

"The only appropriate mechanism for controlling information is classification," says Steven Aftergood, who runs the Project on Government Secrecy for the Federation of American Scientists. "If we want to gain the benefits of university research on problems of national security, we need to conduct it openly. Imposing restrictions short of classification is a slippery slope that will ultimately paralyze the academic process."

Universities have traditionally drawn a sharp line between classified and unclassified information, refusing to accept the ill-defined SBU category. Yet, in a 28 March meeting at the U.S. National Academies, DHS officials and directors of the six university centers funded by the agency discussed draft guidelines to control the dissemination of sensitive information generated by their research. The guidelines were developed by the center directors in collaboration with DHS officials. The academies agreed to be host because of their ongoing interest in the topic.

Besides recommending the scrubbing of papers before publication, the guidelines would have center directors decide whether proposed research projects are likely to produce sensitive information—loosely defined as information not easily available from public sources and/or of potential use to terrorists. Projects that fit that description would be subject to additional scrutiny. The results, says the document, could include "producing different version(s) of the findings for 'For Official Use Only' and for public dissemination, declin[ing] the proposed work, or mov[ing] it to a classified environment."

The guidelines simply acknowledge "the reality of a changing world," says Melvin Bernstein, acting director of DHS's Office of Research and Development, which helped set up the university centers with 3-year renewable grants. "There's an increasing recognition in the university community that there could be circumstances when researchers need to be careful about what can be disseminated."

Although Bernstein says it's too early to know whether the guidelines will become official policy, they appear consistent with a presidential directive issued last December ordering common standards across the

government by the end of 2006 for handling SBU information. After talking with DHS officials, the center directors decided that writing some of the rules themselves would be better than having the government impose them. "We knew we had no choice. This thing was coming our way sooner or later," says Gary LaFree, co-director of the Study of Terrorism and Responses to Terrorism at the University of Maryland, College Park.

One reason that universities have resisted the SBU concept is its vagueness, which some academics fear could lead to federal agencies trying to set arbitrary restrictions on campus research. The executive branch itself seems confused about what information should be withheld from the public and why: The Government Accountability Office reported in March that agencies use 56 different SBU categories in deciding how to control information. Last week, Thomas E. "Ted" McNamara, an official in the Office of the Director of National Intelligence who is leading a federal effort to sort out the confusion, told a congressional panel that some of the government's procedures for handling SBU information "are not only inconsistent but are contradictory." McNamara expects to submit his recommendations next month on standardizing SBU procedures.

But a clearer definition of SBU is unlikely to end the debate. LaFree says the guidelines discussed at the academies meeting could have serious implications for research at the DHS centers. "They could lead to restrictions on the involvement of foreign students and researchers in certain projects," he says, adding that not all center directors are comfortable with the guidelines, despite their role in writing them. "That would be simply unacceptable."

LaFree's concern is not unfounded. In fact, the USC center has been developing procedures—not included in the draft guidelines—that would require foreign nationals to agree to certain conditions before being given access to sensitive information. (Von Winterfeldt won't say what those conditions might be.) Such procedures, critics say, could encourage principal investigators to drop foreigners from sensitive projects. That's already happened in some cases: Yacov Haimes of the University of Virginia in Charlottesville says he deliberately avoided including any foreign nationals when his research team did an unclassified study for the federal government 2 years ago on the risk of a high-altitude electromagnetic pulse attack on the United States.

> "While we have identified several additional effective countermeasures, **only limited details can be revealed for security reasons."**
>
> —Heather Rossof and Detlof von Winterfeldt

That approach could backfire on universities, warns Robert Hardy of the nonprofit Council on Governmental Relations in Washington, D.C. By placing restrictions on publishing, he says, the centers could risk losing the privileges that universities enjoy because they do fundamental research—defined as work whose results are "published and shared broadly within the scientific community." One important privilege is being able to involve foreign nationals in any research project without obtaining a government license.

Randolph Hall, vice president for research advancement at USC and a researcher at the USC center, disagrees with Hardy's interpretation of what qualifies as open publishing. Taking some information out of a paper is not the same as preventing a researcher from publishing, he says, and shouldn't have any bearing on the exemption given to institutions. "It's not unusual for reports at any institution to go through editing, even if some of the changes might be purely grammatical," Hall says. "Similarly, editing out sensitive data is more of a revision than a restriction."

Shaun Kennedy, a chemical engineer and deputy director of the National Center for Food Protection and Defense at the University of Minnesota, Twin Cities, says the proposed guidelines bump up against state laws meant to ensure public access to information. "If I have a For Official Use Only version of a paper in a folder, shredding it would be a violation of the Minnesota Data Practices Act," says Kennedy, adding that the center decided not to start a proposed project analyzing chinks in the nation's food supply chain partly because of that provision. (Instead, the Food and Drug Administration is doing the research internally.)

Some scientists say that there's a more fundamental issue at stake, namely, whether a limit on what goes into the open literature might actually weaken the nation's security. "If you don't publish the information, it might reduce the chances of an attack. But just as likely it could reduce the chances of another researcher coming up with a solution. If the risks are so great, then why shouldn't the research be classified?" asks Toby Smith of the Association of American Universities.

LaFree thinks the argument makes sense. What universities bring to the stable, he says, "is the best minds to look at the data that we pass around. If we end up putting a lot of fences around information, that'll defeat the purpose of doing this type of research in an academic environment."

Von Winterfeldt doesn't believe that a little secrecy will doom research, but he does agree that universities should set and implement policies to protect SBU information. Panels similar to Institutional Review Boards could be set up to do the job, he suggests. And he acknowledges that the panels will have to wrestle with some tough questions. Asked why a sentence in his team's paper on using a helicopter to disperse a dirty bomb didn't qualify as sensitive information, von Winterfeldt said, "It's in the gray zone. I'll discuss it at my next meeting with the author and our staff."

–YUDHIJIT BHATTACHARJEE

Playing it safe. USC researchers removed some details from their paper on the risk and impact of a dirty bomb attack on Los Angeles harbor (*above*) to avoid helping terrorists. Inset shows a model of how radiation might spread.

CREDITS: DAVID MCNEW/GETTY IMAGE; (INSET) ADAPTED FROM A REPORT BY H. ROSOFF AND D. VON WINTERFELDT (OCTOBER 2005)

Published by AAAS

EDITED BY YUDHIJIT BHATTACHARJEE

Celebrities

STUNT SCIENCE. Murat Gunel wasn't just one of the gawkers looking on as performance artist David Blaine spent a week inside a water tank in New York City earlier this month. The Yale neurosurgeon and molecular geneticist headed the medical team that monitored Blaine throughout the stunt, which ended after Blaine came nearly 2 minutes short of setting a record for holding one's breath under water (his time was 7:08). The attempt led Gunel (right) to wonder whether some individuals have genetic quirks that might give them an advantage.

Gunel plans to analyze blood samples from Blaine and from the free divers who helped rescue him after he blacked out. He says he tried to dissuade Blaine, a personal friend, from doing the stunt.

INSIDE GOVERNMENT

NEW USGS HEAD. A principled resignation has proved lucky for petroleum geologist Mark Myers, who last week was nominated as the next director of the U.S. Geological Survey (USGS). Environmentalists like the fact that Myers, 51, left as director of Alaska's Division of Oil and Gas last fall along with five other officials because he felt that a pipeline deal negotiated by Governor Frank Murkowski with several oil companies would shortchange the state. "He has a significant amount of integrity," says Karen Wayland of the Natural Resources Defense Council in Washington, D.C.

Myers holds a Ph.D. in sedimentology from the University of Alaska and worked in industry before joining the state agency, which leases drilling rights to oil and gas companies. He also headed the state's geological survey. Robert Swenson, acting state geologist, says Myers "first and foremost is a broad-based scientist. He's fair, and he stands up for the people who work for him."

Myers, who relied on information from USGS to make decisions in his previous job, says one of his major goals will be to ensure that data produced by the agency "remains objective." If confirmed by the Senate, he will replace Charles Groat, who resigned in June 2005 after nearly 7 years as director to return to academia.

Got a tip for this page? E-mail people@aaas.org

ROTATING CHAIRS. Steven Beering, president emeritus of Purdue University in West Lafayette, Indiana, was all set to head a high-profile look at the state of U.S. science and math education—until the National Science Board, which created the education commission, realized that it needed him as its leader.

Warren Washington, its current chair, had completed a 12-year stint on the presidentially appointed board, which oversees the National Science Foundation (NSF), and last week members elected Beering as his successor. But that left a vacancy at the top of the board's new education commission (*Science*, 7 April, p. 45).

That void has been filled by physics Nobelist Leon Lederman, who founded the Illinois Mathematics and Science Academy. Lederman is one of 12 public members of the new panel, which includes former Ohio congressman Louis Stokes as well as a middle school science teacher from Lederman's home state. The commission expects to issue a report next spring.

MONEY MATTERS

WAR ON CANCER. Real estate tycoon and publisher Mortimer Zuckerman has gifted $100 million to the Memorial Sloan-Kettering Cancer Center (MSKCC) in New York City. The center is saying thanks by putting his name on a new 23-story laboratory building scheduled to open this month. Zuckerman, who is on MSKCC's board, says he made the gift "to accelerate the pace of progress" in cancer research and "to help the center's extraordinary scientists and physicians achieve their crucial goals."

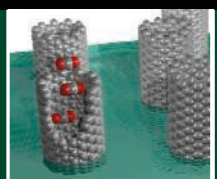
Movers >>

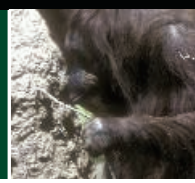
BUSINESS SENSE. John Chisholm, 59, the newly appointed chair of the U.K. agency that oversees government biomedical research, has one big advantage over his predecessor Anthony Cleaver: He actually studied science at university. Chisholm earned a degree in mechanical sciences from the University of Cambridge and joined a computing arm of British Petroleum before helping launch a successful U.K. software company called CAP Scientific. His signature accomplishment may have been forging a disparate group of U.K. military labs into a single research unit, later dubbed QinetiQ. It was spun off as a private firm in 2001, and a stock sale this year raised more than $1 billion.

Those corporate management skills will come in handy at the Medical Research Council (MRC) as Chisholm follows orders to merge its research with clinical studies in the Department of Health. Some scientists worry that basic science could be hurt in the shuffle. Chisholm offers a reassuring view: "I have long been a passionate advocate for research. … It's a wonderful moment to have been given a chance to contribute to seizing the opportunities before the MRC."

CREDITS (TOP TO BOTTOM): BRENDAN MCDERMID/REUTERS; KAYA BILGUVAR; FORREST CRANE/*PETROLEUM NEWS*

Published by AAAS





LETTERS

edited by Etta Kavanagh

Translation Research and Drug Development

JOCELYN KAISER'S RECENT ARTICLE ON "TRANSLATIONAL research" ("A cure for medicine's ailments?," News Focus, 31 Mar., p. 1852) sounded an encouraging note to basic and clinical researchers alike who yearn to test their pet ideas for new cures. According to Kaiser, translational research is loosely defined as "moving a basic discovery into early clinical trials." However, NIH's apparent desire to foster translational research by funding university-based drug development centers sends shivers down this taxpayer's back. Pharma spends upwards of $800 million and takes 10 to 12 years to get a drug from bench to bedside (*1*). Annual R&D investment by pharma has risen from $1 billion to $40 billion since 1975, while annual new drug approvals have remained flat at between 20 and 30. Thus, drug development today is less efficient than 30 years ago, which partly explains the continual rise in drug costs. Although NIH's interest in drug development is laudable, does anybody truly believe that academic translational research centers will be as efficient, let alone competitive, at developing drugs as pharma?

Kaiser pointed to an anecdotal case where a single-minded researcher persevered for years to get a novel anticancer agent tested in a small clinical trial. The implication was that the researcher could have made more rapid progress if her university had invested in more translational research activities. Even if this were true, who will fund the rest of the costly activities required to bring this drug to market? Granted, these activities may fall outside of the accepted view of translational research. But without a funding partner, investing in translational research is akin to building a bridge to nowhere.

The road from the discovery of a drug to the first human clinical trial leads through a painstaking and circuitous route that is tedious and expensive, fails more than 90% of the time, does not lead to front-line publications, and does not constitute the type of research that many deem worthy of a Ph.D. But it will make or break your favorite drug candidate. I believe that a better use of taxpayers' dollars would be to support innovative research proposals related to improving the efficiency of the drug R&D process. In this way, we will lower the time and cost, as well as the failure rate, of bringing new drugs to market, and the public will benefit. And I bet pharma will invest private dollars into these activities. This is the sort of translational research that makes more sense to me—building bridges between academia and pharma—than trying to duplicate pharma activities in academic settings.

JOHN ERICKSON

Sequoia Pharmaceuticals, Inc., 401 Professional Drive, Gaithersburg, MD 20879, USA. E-mail: john.erickson@sequoiapharma.com

Reference

1. J. A. DiMasi, R. W. Hansen, H. G. Grabowski, *J. Health Econ.* **22**, 151 (2003).

Extinct or Possibly Extinct?

LISTS OF EXTINCT SPECIES OFTEN ACT AS "WAKE-UP calls" and are based on the length of time since the last sighting, resulting in numerous species having been prematurely classified as being extinct only to be rediscovered (*1*). This not only provides ammunition for environmental sceptics (D. S. Wilcove, "Rediscovery of the ivory-billed woodpecker," Perspectives, 3 June 2005, p. 1422) but also undermines potential conservation action and, more worryingly, public support (*2*). It is almost impossible to determine with any certainty whether a species is extinct. Therefore, any statement of extinction is probabilistic by nature (*3*). The rediscovery of the ivory-billed woodpecker [J. W. Fitzpatrick *et al.*, "Ivory-billed woodpecker (*Campephilus principalis*) persists in continental North America," Reports, 3 June 2005, p. 1460] has recently been called into question [(*4*); D. A. Sibley *et al.*, Comment on "Ivory-billed woodpecker (*Campephilus principalis*) persists in continental North America," Technical Comment, 17 Mar., www.sciencemag.org/cgi/content/full/311/5767/1555a]. Even so, it raises the question, which seems to have been missed by scientists, as to whether this species should have been declared extinct in the first place.

The case for classifying the ivory-billed woodpecker as extinct was based on the very long time that had elapsed since the most recent confirmed sighting. Under the IUCN Red List criteria, a species is classified as "extinct" only after exhaustive surveys fail to produce any observations over an appropriate time period and geographical range (*5*). For most species, this is impractical (*2*).

A statistical test for extinction based on the most recent sightings of a species was described by Solow (*6*). If we use the five most recent pre-2004 sightings of the ivory-billed woodpecker (1938, 1939, 1941, 1944, and 1952) (*7*), then the significance level (or *P* value) in testing in 2004 for extinction is 0.186. The hypothesis that the ivory-billed woodpecker is extant should not have been rejected. Even if we take the last sighting to be 1944, as others suggest (*4*), then the significance level is 0.056. This raises the question of whether the IUCN Red List requires a "possibly extinct" category as any statement of extinction is probabilistic by nature.

DAVID L. ROBERTS

Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AB, UK. E-mail: d.roberts@kew.org

References

1. S. Pimm, *Nature* **426**, 235 (2003).
2. G. J. McInerny *et al.*, *Conserv. Biol.* **20,** 562 (2006).
3. D. L. Roberts, A. C. Kitchener, *Biol. Conserv.* **128**, 285 (2006).

4. J. A. Jackson, *Auk* **123**, 1 (2006).
5. IUCN, *IUCN Red List Categories and Criteria: Version 3.1.* (IUCN, Gland, Switzerland and Cambridge, UK, 2001).
6. A. J. Solow, *Math. Biosci.* **195**, 47 (2005).
7. E. Fuller, *Extinct Birds* (Cornell Univ. Press, Ithaca, NY, 2001).

Incorporating Evolution into Medical Education

IN THEIR EDITORIAL "MEDICINE NEEDS EVOLUTION" (24 Feb., p. 1071), R. M. Nesse *et al.* highlight human maladies whose origin and expression might be illuminated by evolutionary perspectives. The examples are many, and they point out the need for a central evolutionary insight that can help to inform all of medical thinking and serve as the basis for the integration of evolution into medical education and clinical practice.

Medicine might benefit most from embracing evolution theory's recognition of individual variation within populations of organisms, a property that Ernst Mayr has called "the cornerstone of Darwin's theory of natural selection" (*1*). This "population thinking," as Mayr calls it, helped to undo typological thinking in biology, and it can help to dismantle typological notions of disease by highlighting individual differences in disease susceptibility and expression, as well as variations in response to treatment.

The inextricable relationship between evolution and genetics is evident in current genomic-based efforts such as the HapMap project, which catalogs DNA variants associated with disease, and in the recently announced Genes and Environment Initiative at NIH, which will investigate the interaction of genetic and environmental variations in common diseases. A major challenge for medical education is to incorporate genetics and evolution into education systems where neither receives the attention necessary to make it a routine part of medical thinking or clinical practice.

JOSEPH D. MCINERNEY

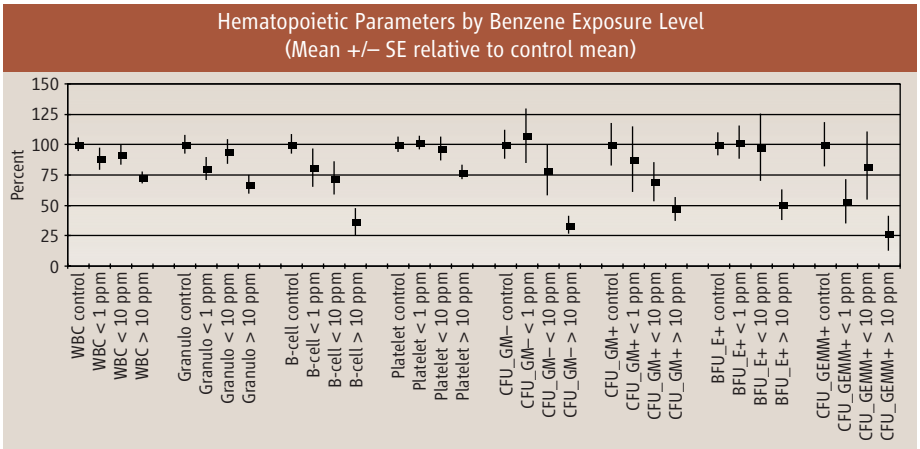
Executive Director, National Coalition for Health Professional Education in Genetics, 2360 West Joppa Road, Suite 320, Lutherville, MD 21093, USA.

Reference

1. E. Mayr, *One Long Argument: Charles Darwin and the Genesis of Modern Evolutionary Thought* (Harvard Univ. Press, Cambridge, MA, 1991).

Benzene Exposure and Hematotoxicity

IN THEIR REPORT "HEMATOTOXICITY IN WORKERS exposed to low levels of benzene" (3 Dec. 2004, p. 1774), Q. Lan *et al.* present data on blood cell counts and hematopoietic progenitor cell colony formation from sera of benzene-exposed workers (and controls) in China, from which they conclude that their data demonstrate hematotoxicity with benzene air levels at less than 1 ppm. Although we concur that their data demonstrate hematotoxicity with benzene levels at greater than 10 ppm, we do not observe in their data consistent evidence of hematotoxicity at lower levels.



Hematopoetic parameters by benzene exposure level. Mean (±SE) relative to control mean. Adapted from Lan *et al.*

Their blood cell counts (their table 1) showed a monotonically increasing effect only for platelets and B cells, but not for the measured cell lines that might be expected to lead to myeloid leukemic lines. White blood cell and granulocyte counts that showed a reduction in cell number at less than 1 ppm did not show a further reduction among workers with exposures up to 10 ppm.

The authors' progenitor cell colony formation data (their fig. 1) did not separate out the data below 10 ppm and thus do not demonstrate whether an effect occurred at <1 ppm. They have kindly supplied us those data (our figure). In these data, we observe a suggestive monotonically increasing trend only for granulocyte-macrophage colony-formation (CFU_GM–), which first appears at greater than 1 ppm in the absence of erythropoietin and at less then 1 ppm in the presence of erythropoietin. Neither reduction is statistically significant until the group with benzene exposure at greater than 10 ppm is considered.

We consider the authors' conclusion premature, based only on the difference of reduction in in vitro granulocyte-macrophage colony formation by the addition of erythropoietin to the culture medium. The only implication of the difference of adding erythropoietin is that by driving the formation of the erythroid lineage, they reduce the myeloid colony numbers ("lineage competition").

A demonstration of damage to stem cell function or number would be a more relevant indication of hematotoxicity than is damage to committed progenitor stem cells as proposed by Lan *et al.* We would propose the alternative conclusion that their data show that hematotoxicity as measured by reduction of in vitro colony formation may well be ascribed to levels of benzene greater than 10 ppm but do not justify the implied damage from levels lower than that.

Finally, although the authors' findings of reduction in peripheral granulocytes may carry statistical significance, the numbers they found in their exposed individuals are all fully within the normal range and do not carry clinical significance.

STEVEN H. LAMM[1] AND HANS W. GRÜNWALD[2]

[1]Consultants in Epidemiology & Occupational Health, LLC, 3401 38th Street, NW, #615, Washington, DC 20016, USA, and Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD 21205, USA. [2]Division of Hematology-Oncology, Queens Hospital Center–Cancer Center, Jamaica, NY 11432, USA, and Department of Medicine, Mt. Sinai School of Medicine, New York, NY 10029, USA.

Response

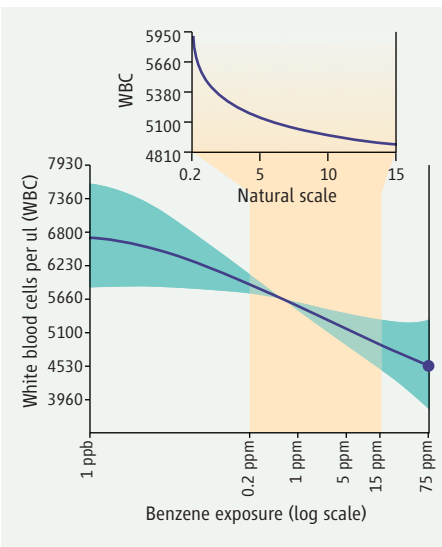
WE REPORTED THAT WHITE BLOOD CELL (WBC) counts were decreased in workers exposed to less than 1 ppm benzene compared with controls and that a highly significant dose-response relationship was present (original Table 1, text). Lamm and Grünwald argue that a monotonic dose-response relationship must be present across higher levels of exposure before one can accept differences between controls and the lowest exposure group. Although we do not necessarily agree with their premise, we confirmed the monotonicity of the association by spline regression analyses of WBC count and benzene exposure and found no apparent threshold within the occu-

pational exposure range of our study (0.2 to 75 ppm benzene; see our figure).

Another goal of our study was to determine whether benzene was associated with a decrease in progenitor cells across a wide range of exposure, and whether progenitor cells were more sensitive to the effects of benzene than mature cells. We found highly statistically significant, inverse dose-dependent trends for all progenitor cells and observed that a number of progenitor cells, including CFU-GEMM colonies, were significantly more sensitive to the effects of benzene than peripheral WBC or granulocyte counts among highly exposed workers (original fig. 1). Lamm and Grünwald show progenitor colony data for each exposure category and break out effects for the <1 ppm group, even though there are only 8 subjects in this category (their figure), and report that colony counts in this group were not significantly different from controls. A substantially larger study would be needed to address this question, which was not a goal of our paper. In addition, they present data on WBC and other cell counts for the subgroup of 53 subjects with progenitor colony data (their figure); conclusions based on benzene exposure and mature blood cell counts should be based on the entire data set (original Table 1; our figure) rather than on this subgroup.

Lamm and Grünwald suggest that it would have been worthwhile to culture stem cells. Although data of this type would be of interest, it was not feasible to collect in the occupational setting, and CFU-GEMM, CFU-GM, and BFU-E are commonly used surrogates for stem cell measurements.

Finally, we note that changes of the magnitude we report for mature blood cells are generally considered unlikely to have immediate clinical consequences. However, as we show even more pronounced effects in progenitor cells, there is a concern that the overall pattern of hematologic changes we observe could reflect events in bone marrow that may be associated with health effects in the future, particularly among genetically susceptible subpopulations (*1*–*3*).



Plot shows the dose-response curve (line) and 95% pointwise confidence limits (shaded areas) for differences between white blood cell (WBC) count at a given air benzene exposure versus WBC level at a reference dose of 0.6 ppm (median benzene exposure level of the total study population). Graph shows the fitted nonparametric response curve using generalized additive models (*4*) on a natural scale versus benzene exposure on a log scale (truncated at 1 ppb); inset graph shows the fitted nonparametric response curve on a natural scale versus benzene exposure between 0.2 and 15 ppm on a natural scale. The nonparametric curve was fitted using a regression spline with 1 segment, which was the optimal number of polynomial segments (1 to 5 tested) based on the Akaike Information Criterion (*5*). The model was adjusted for the same variables used in previous analyses (original table 1). Complete data from 139 controls and 247 exposed subjects were available. Data were used from only the first study year for subjects with repeat measures in the second study year. Using data from only the second year for these subjects resulted in essentially the same prediction models. Air benzene exposure among the controls was estimated based on the linear relation of log urinary benzene levels on log air benzene (*6*). The slope of the spline function was significantly less than zero for every point between 0.2 and 15 ppm, indicating that the geometric mean WBC count decreased significantly with increasing exposure over this specific exposure range ($P < 0.05$, accounting for multiple comparisons).

QING LAN,[1] ROEL VERMEULEN,[1] LUOPING ZHANG,[2] GUILAN LI,[3] PHILIP S. ROSENBERG[1], BLANCHE P. ALTER,[1] MIN SHEN,[1] STEPHEN M. RAPPAPORT,[4] RONA S. WEINBERG,[5] STEPHEN CHANOCK,[1,6] SURAMYA WAIDYANATHA,[4] CHARLES RABKIN,[1] RICHARD B. HAYES,[1] MARTHA LINET,[1] SUNGKYOON KIM,[4] SONGNIAN YIN,[3] NATHANIEL ROTHMAN,[1] MARTYN T. SMITH[2]

[1]Division of Cancer Epidemiology and Genetics, National Cancer Institute (NCI), National Institutes of Health (NIH), Department of Health and Human Services (DHHS), Bethesda, MD 20892, USA. [2]School of Public Health, University of California, Berkeley, Berkeley, CA 94720, USA. [3]Chinese Center for Disease Control and Prevention, Beijing, China. [4]School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. [5]New York Blood Center, Clinical Services, White Plains, NY 10605, USA. [6]Center for Cancer Research, NCI, NIH, DHHS, Bethesda, MD 20892, USA.

References

1. T. Hastie *et al.*, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer-Verlag, Berlin, 2002).
2. H. Akaike, in *Second International Symposium on Information Theory*, B. N. Petrov, F. Csàki, Eds. (Akademia kiadó, Budapest, 1973), pp. 267–281.
3. S. Kim *et al.*, *Carcinogenesis*, 8 Dec. 2005; Epub ahead of print.
4. S. N. Yin *et al.*, *Br. J. Ind. Med.* **44**, 124 (1987).
5. N. Rothman *et al.*, *Cancer Res.* **57**, 2839 (1997).
6. Q. Lan *et al.*, *Cancer Res.* **65**, 9574 (2005).

CORRECTIONS AND CLARIFICATIONS

NetWatch: "All physics, all the time" (28 Apr., p. 505). The item incorrectly stated that Bowling Green State University is in Kentucky. It is in Ohio.

ScienceScope: "NYU gift kicks up more dust" by M. Balter (28 Apr., p. 513). The URL for the "Statement of Concern" mentioned in the item was incorrect. It should be www.bib-arch.org/bswbOOunprovenanced.html.

News of the Week: "Opening the door to a chilly new climate regime" by R. A. Kerr (21 Apr., p. 350). The current abbreviated "ACC" was incorrectly identified. It is the Antarctic Circumpolar Current.

Special Section News: "A one-size-fits-all flu vaccine?" by J. Kaiser (21 Apr., p. 380). The table is missing a symbol indicating that "DNA vaccine with NP, sometimes M2 genes" stimulates cytotoxic T lymphocytes.

TECHNICAL COMMENT ABSTRACTS

COMMENT ON "The Brain of LB1, *Homo floresiensis*"

R. D. Martin, A. M. MacLarnon, J. L. Phillips, L. Dussubieux, P. R. Williams, W. B. Dobyns

Endocast analysis of the brain *Homo floresiensis* by Falk *et al*. (Reports, 8 April 2005, p. 242) implies that the hominid is an insular dwarf derived from *H. erectus*, but its tiny cranial capacity cannot result from normal dwarfing. Consideration of more appropriate microcephalic syndromes and specimens supports the hypothesis of modern human microcephaly.

Full text at www.sciencemag.org/cgi/content/full/312/5776/999b

RESPONSE TO COMMENT ON "The Brain of LB1, *Homo floresiensis*"

Dean Falk, Charles Hildebolt, Kirk Smith, M. J. Morwood, Thomas Sutikna, Jatmiko, E. Wayhu Saptomo, Barry Brunsden, Fred Prior

Martin *et al*. claim that they have two endocasts from microcephalics that appear similar to that of LB1, *Homo floresiensis*. However, the line drawings they present as evidence lack details about the transverse sinuses, cerebellum, and cerebral poles. Comparative measurements, actual photographs, and sketches that identify key features are needed to draw meaningful conclusions about Martin *et al.*'s assertions.

Full text at www.sciencemag.org/cgi/content/full/312/5776/999c

RISK ASSESSMENT

Predicting Asbestos's Fallout

Jonathan M. Samet

A century ago, asbestos seemed a material ideally matched to the needs of increasingly industrialized and motorized Western societies. It has the useful properties of heat- , fire-, and chemical resistance along with strength and flexibility (*1*). Consequently, it became widely used in building materials, friction products, and fire-retarding fabrics. But after peaking in the 1970s, asbestos consumption fell quickly as recognition of the risks it posed to health led to bans and substitution with other materials. Although sentinel cases of asbestosis (the scarring disorder of the lungs caused by inhaling asbestos fibers) were reported as early as 1900, asbestos was not widely recognized as causing cancer until the 1950s and 1960s, when epidemiological and clinical studies linked asbestos exposure to mesothelioma (cancer of the lining of the lung and abdominal cavity) and other lung cancers. The identification of asbestos as a carcinogen was delayed because the increased risks for these cancers only become apparent decades after first exposure.

Forecasting Product Liability Claims
Epidemiology and Modeling in the Manville Asbestos Case

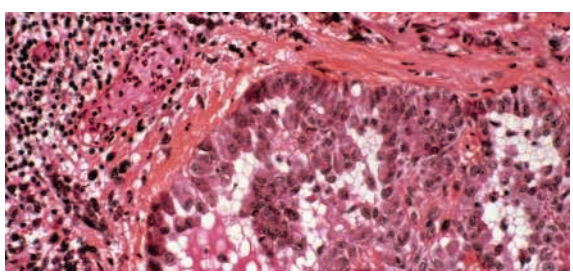
by Eric Stallard, Kenneth G. Manton, and Joel E. Cohen

Springer, Berlin, 2005. 424 pp. $84.95, £69, €89.95. ISBN 0-387-94987-9. Statistics for Biology and Health.

Millions of workers in the United States and other countries have now been exposed to asbestos. Many have developed asbestos-caused diseases, and millions of current and former workers are still at risk. Thousands of lawsuits on behalf of affected workers have been filed against companies that processed asbestos and made asbestos-containing products. The costs of compensating the claims led to bankruptcy for many companies, including the Johns-Manville Corporation, which in 1982 filed for Chapter 11 bankruptcy on the basis of the numbers of claims already filed or anticipated. The numbers of lawsuits and the costs to industry and insurers prompted calls in the United States for a federal legislative remedy, but attempts to pass such legislation have been unsuccessful to date.

Predicting the future course of litigation and planning legal remedies require estimates of the future numbers of claims for asbestos-caused disease. How can we estimate the burden of disease caused in a population by asbestos (or any other avoidable risk factor)? In a thoughtful 1953 paper, the epidemiologist Morton Levin proposed a formula for this calculation that is now widely referred to as the attributable risk (*2*). In his formula, the proportion of disease attributable to a factor increases as the frequency of the exposure increases and as the disease risk associated with exposure increases. Even with the data available in 1953, the formula provided an estimate indicating that the majority of lung cancer cases in males could be attributed to smoking.

The concept of attributable risk has been extended in contemporary quantitative risk assessment, a formalism codified in a 1983 report of the U.S. National Research Council (*3*). Such risk assessments first involve the determination that an exposure poses a risk, "hazard identification," and, if appropriate, a characterization of population risk ("risk characterization"). Consistent with Levin's formula, estimating the risk requires information on the exposure of the population ("exposure assessment") and on the effect at each level of exposure ("dose-response"). In most instances, risk assessment requires assumptions, but these can be systematically tabulated and consideration given to the consequences of the resulting uncertainties. Quantitative risk assessment is a tool widely used in decision-making and legally required for some classes of environmental agents.



Mesothelioma cells.

In *Forecasting Product Liability Claims*, demographers Eric Stallard (Duke University), Kenneth Manton (Duke University), and Joel Cohen (Rockefeller University) use a risk assessment framework to estimate the numbers of claims expected during the period between 1990 and 2049 for asbestos-related disease among men exposed to Johns-Mansville–produced asbestos and asbestos products. This work came from their membership on an expert panel appointed—under Rule 706 of the Federal Rules of Evidence—by Jack Weinstein, the U.S. District Court judge who was administering the Manville Personal Injury Settlement Trust. When he was assigned jurisdiction over the Trust in 1990, it was failing, tens of thousands of claims had been filed against it, and more were anticipated. To ensure that sufficient funds would be available for payouts long into the future, Weinstein needed to be able to anticipate the number and nature of claims that would be made against the Trust. This monograph sets out in full detail the quantitative forecasting model that the expert panel developed to help Weinstein resolve the Johns-Manville litigation. In its foreword, the judge offers sharp insights into asbestos litigation and the need for legislative remedy.

The authors' work was not the first risk assessment applied to workers exposed to asbestos. A 1982 report by Alexander Walker offered projections of future disease in support of the decision by Johns-Manville to file for Chapter 11 bankruptcy (*4*). Irving Selikoff, whose pioneering epidemiological studies and sustained advocacy brought the epidemic of asbestos-caused disease into the spotlight, had also provided forecasts.

Stallard, Manton, and Cohen faced the challenge of estimating the two components of Levin's formula: the number of exposed individuals and the risk of their becoming ill and making claims. The authors begin with a detailed and critical review of the earlier models. Not surprisingly, they found flaws, but some of the same underlying conceptual approaches are incorporated in their model as well. They estimated the exposed population by using national data on mesothelioma occurrence to project the size of the source population that yielded the observed cases, relying on a model for the relation between asbestos exposure and mesothelioma occurrence. The annual numbers of claims that would be made for the estimated exposed population were based on the experience of the Manville Trust over the interval between 1990 and 1992. In the text, the authors set out their assumptions and computations in detail and provide extensive sensitivity analyses. The authors' model yields a staggering estimate of cumulative claims through the middle of the next century—517,000—and thousands of claims are projected annually from now until 2049.

Because estimates of numbers of claims will be needed to develop a federal compensation system, the book's model could underlie the implementation of such a plan. Johns-Manville Corporation, although the largest asbestos com-

The reviewer is at the Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD 21205, USA. E-mail: jsamet@jhsph.edu. He is now an editor of the Springer series in which the book appears but was not involved in the project.

pany operating in the United States, did not have the majority of the asbestos market. The authors' calculations, which can be reasonably generalized to other companies, imply that hundreds of thousands of claims may still be filed in the United States before the epidemic of asbestos-caused disease comes to its end.

Forecasting Product Liability Claims is notable for its illustration of the possibility of using epidemiologic and demographic methods to develop models for broad policy purposes. It also documents a successful instance of asking court-appointed experts to provide guidance on a highly adversarial issue. Nonetheless, only those interested in the details of the models will want to read the book from cover to cover. It is formula-rich and dense in its description of data sources and the machinery of the models, as it should be. Readers with interests in environmental or occupational health, product liability, or science and the law may prefer to scan the book to gain an appreciation of the approach.

References

1. R. Maines, *Asbestos and Fire: Technological Trade-offs and the Body at Risk* (Rutgers Univ. Press, New Brunswick, NJ, 2005).
2. M. L. Levin, *Acta Unio Int. Contra Cancrum* **9**, 531 (1953).
3. National Research Council, *Risk Assessment in the Federal Government: Managing the Process* (National Academy Press, Washington, DC, 1983); www.nap.edu/books/0309033497/html.
4. A. M. Walker, *Projections of Asbestos-Related Disease 1980–2009* (Epidemiology Resources, Chestnut Hill, MA, 1982).

10.1126/science.1127406

CLIMATE CHANGE

Weather Effects

Andrew Goudie

The Winds of Change
Climate, Weather, and the Destruction of Civilizations

by Eugene Linden

Simon and Schuster, New York, 2006. 314 pp. $26, C$36. ISBN 0-684-86352-9.

Ellsworth Huntington, a lapsed missionary who traveled in Central Asia early in the 20th century and subsequently became a celebrated professor of geography at Yale, was intrigued by the evidence for the rise and fall of civilizations that he encountered on his expeditions. This led him to write *The Pulse of Asia* (*1*), a 1907 book in which he argued that climatic changes had been a major cause of great events in human history. During his career he developed this theme for other regions, but his work became the subject of considerable criticism. In part, his critics objected to his views on the ways in which climate influenced racial characteristics and the energies of different peoples: bracing climates produced energetic people like the British or the Icelanders, while hot, sultry climates produced more indolent characteristics. That said, Huntington's historical studies, although largely forgotten, are original in conception and were based on paleoclimatic research that was, for the time, of a high quality. Unfortunately, the means of dating and paleoenvironmental reconstruction available to Huntington were insufficiently precise.

In recent decades, however, the situation has changed because of the development, from such sources as cores drilled through the polar ice caps, of high-resolution chronologies for climate change. We now have much clearer pictures of the abruptness of some environmental changes and the degree of climate change that has taken place in what had once been regarded as the stable Holocene (i.e., the last circa 10,000 years). These data have led to a renewed interest in the idea that Huntington developed, that the rise and fall of civilizations can owe much to abrupt severe climatic change. The archaeologist Brian Fagan has recently produced books on this theme (*2*, *3*), and now environmental journalist Eugene Linden has done the same.

In *The Winds of Change*, Linden contends that it is not just our new appreciation of the nature of climate change in the Holocene that has revolutionized thought. There is also the realization that, perhaps counterintuitively, more advanced societies are less able to respond to climatic stresses than mobile bands of hunter-gatherers. In spite of our sophistication, new technologies, and degree of social organization, we have, he argues, become more vulnerable to even smaller disruptions of climate. For example, a farmer living in the Arctic may not survive if the growing season becomes too short and the harbors become iced up. Likewise, a city dependent on irrigated agriculture may fail if the water supply dries up. However, it is not just that the direct effects of climatic change (such as floods and droughts) may be fateful; one must also consider such indirect consequences as disease, blight, and civil disorder.

So in the first part of the book, Linden explores the evidence that the cold snap around 8200 years ago set back the growth of complex societies in the Levant, drought killed off the Akkadian civilization 4200 years ago, the collapse of the Mayans 1100 years ago was climate-related, and the Late Medieval abandonment of the Greenland settlements was caused by the onset of the Little Ice Age. Next, he examines the evidence that climate has changed during the Holocene, analyzing the nature of shifts in the oceans' large-scale thermohaline circulation and the information retrieved from ice and ocean sediment cores.

Frozen out. A wedding at Hvalsey Church in 1408 provided the last historical account from the Norse settlements in Greenland.

CREDIT: WOLFGANG KAEHLER/CORBIS

Linden then revisits the possible explanations for the calamities and events discussed in the first part, assessing climate in relation to other hypotheses such as political upheaval and environmental degradation. He also examines the nature, causes, and consequences of recent El Niño events—including the roles they may have played in British colonial domination in India, the French Revolution, the demise of the Suharto regime in Indonesia, and the fall of Ethiopia's Emperor Haile Selassie. Lastly, Linden raises the uncomfortable question "Are we next?" and points to the accelerating pace of climate change over the last five decades.

Beautifully written, *The Winds of Change* is a very thought provoking volume. Linden manages to weave history, science, and narrative together in a compelling way. It is a shame that those who are stimulated to delve deeper will find the text for the most part lacks references. Nonetheless, Ellsworth Huntington would have liked the book hugely.

References

1. E. Huntington, *The Pulse of Asia: A Journey in Central Asia Illustrating the Geographic Basis of History* (Houghton Mifflin, Boston, 1907).
2. B. M. Fagan, *Floods, Famines, and Emperors: El Niño and the Fate of Civilizations* (Basic, New York, 1999).
3. B. M. Fagan, *The Long Summer: How Climate Changed Civilization* (Basic, New York, 2004).

10.1126/science.1128689

The reviewer is at St. Cross College, St. Giles, Oxford OX1 3LZ, UK. E-mail: andrew.goudie@st-cross.oxford.ac.uk

ENVIRONMENT

Investments Toward Sustainable Development

Environmental goals cannot be attained without also addressing poverty; similarly, addressing poverty is essential for improving the environment; both need additional resources, particularly in developing nations.

Jeffrey D. Sachs and Walter V. Reid*,†

Sustainable development, meaning economic growth that is environmentally sound, is a practical necessity. Environmental goals cannot be achieved without development. Poor people will circumvent environmental restrictions in their desperation for land, food, and sustenance. Nor can development goals be achieved and sustained without sound environmental management. Environmental catastrophes will undermine economic life, whether in New Orleans or Nigeria. Therefore, investing in poverty reduction is crucial for environmental policy, while investing in the environment is vital for successful poverty reduction (see figure, right). Yet the world underinvests in both, and rich-country and poor-country governments overlook the policy links between poverty reduction and the environment.

The United Nations (U.N.) Millennium Project (*1*) and the Millennium Ecosystem Assessment (MA) (*2*) highlighted the centrality of environmental management for poverty reduction and general well-being. Each report emphasized the unsustainability of our current trajectory. Millions of people die each year because of their poverty and extreme vulnerability to droughts, crop failure, lack of safe drinking water, and other environmentally related ills. The desperation of the poor and heedlessness of the rich also exact a toll on future well-being in terms of habitat destruction, species extinction, and climate change.

The goal of the Millennium Project (*1*) is to develop and to promote practical plans for achieving the U.N. Millennium Development Goals (MDGs) for ending poverty, eradicating hunger, achieving universal primary education, improving health, and restoring a healthy environment. The MA, in turn, examined the consequences of ecosystem change for human well-being and analyzed options for conserving ecosystems while enhancing their contributions to people (*2, 3*). The MA and the Millennium Project reached strikingly parallel conclusions:

Environmental degradation is a major barrier to the achievement of the MDGs. The MA examined 24 ecosystem services (the benefits people obtain from ecosystems) and found that productivity of only 4 had been enhanced over the last 50 years, whereas 15 (including capture fisheries, water purification, natural hazard regulation, and regional climate regulation) had been degraded. More than 70% of the 1.1 billion poor people surviving on less than $1 per day live in rural areas, where they are directly dependent on ecosystem services.

An investment strategy for sustainable development in low-income countries.

Investing in environmental assets and management are vital to cost-effective and equitable strategies to achieve national goals for relief from poverty, hunger, and disease. For example, investments in improved agricultural practices to reduce water pollution can boost coastal fishing industry. Wetlands protection can meet needs of rural communities while avoiding costs of expensive flood control infrastructure. Yet these investments are often overlooked.

Reaching environmental goals requires progress in eradicating poverty. More coherent and bolder poverty reduction strategies could ease environmental stresses by slowing population growth and enabling the poor to invest long term in their environment.

We recommend the following measures in 2006. First, we call on the rich donor countries to establish a Millennium Ecosystem Fund to give poor countries the wherewithal to incorporate environmental sustainability into national development strategies. The fund would support work that focuses on how poverty reduction can enhance environmental conservation (e.g., by giving farmers alternatives to slash and burn) and how environmental sustainability can support poverty reduction (e.g., watershed management to maintain clean water supplies). It would also support national ecosystem service assessments to help decision-makers factor the economic and health consequences of changes in ecosystem services into their planning choices.

The fund would initially need roughly $200 million over 5 years. It would enable universities and scientists in dozens of the poorest countries to incorporate the science of environmental sustainability into poverty reduction strategies. The programs would generate evidence for countries to use in setting priorities for national development and environmental investments.

Second, the United Nations should establish a cycle of global assessments modeled on the MA and similar to the climate change reports produced at 4- to 5-year intervals by the Intergovernmental Panel on Climate Change (IPCC). The MA and IPCC cost roughly $20 million, and each mobilized in-kind contributions of that magnitude. A global network of respected ecologists, economists, and social scientists working to bring scientific knowledge to decision-makers and to the public can clarify the state of scientific knowledge, help to mobilize needed research, and defeat the obfuscation led by vested interests.

France's recent initiative for a consultative process exploring the merits of an International Mechanism of Scientific Expertise on Biodiversity (*4*) could be one means of establishing a regular assessment process if, along with biodiversity, it also addresses the linkages between ecosystem change and human well-being. Also, it would need to evaluate potential policy, institutional, and behavioral responses.

Third, the world scientific community needs to chart an interdisciplinary strategy for sustainable development research, backed by increased funding. Leading scientific institutions should now coalesce behind a shared agenda on sustainable development and thereby help to draw governments into the challenges of the 21st century.

J. D. Sachs is director of the U.N. Millennium Project and director of the Earth Institute, Columbia University, New York, NY 10027, USA. W. V. Reid was director of the Millennium Ecosystem Assessment and is with the Woods Institute for the Environment, Stanford University, Stanford, CA 94305, USA.

*Present address: David and Lucile Packard Foundation, Los Altos, CA 94022, USA.

†Author for correspondence. E-mail: wreid@packard.org

References

1. UN Millennium Project, *Investing in Development: A Practical Plan to Achieve the Millennium Development Goals* (Earthscan, London and Sterling, VA, 2005).
2. Millennium Ecosystem Assessment, *Ecosystems and Human Well-being: Synthesis* (Island Press, Washington DC, 2005).
3. Millennium Ecosystem Assessment, *Global Assessment Reports,* vol. 1, *Current State and Trends*; vol. 2, *Scenarios*; vol. 3, *Policy Responses;* vol. 4, *Multiscale Assessments* (Island Press, Washington, DC, 2005).
4. International Mechanism of Scientific Expertise on Biodiversity (www.imoseb.net/).

10.1126/science.1124822

MATERIALS SCIENCE

Making High-Flux Membranes with Carbon Nanotubes

Separation of gases and liquids into their constituents is crucial for many industrial processes. Membranes that incorporate nanotubes show promise for high selectivity and throughput.

David S. Sholl and J. Karl Johnson

In 1871, James Clerk Maxwell devised a thought experiment whereby a "demon" separates molecules without performing work. The practical realization of such a demon would be extraordinarily useful, because at present, vast amounts of energy and money are expended around the globe separating chemical mixtures. Even separation of very simple molecules can have enormous implications. For example, purifying O_2 and N_2 from air is a multibillion dollar industry, and the ability to economically separate CO_2 from power plant flue gases could revolutionize efforts to reduce greenhouse gas emissions. Efficient membranes are real-world analogs of Maxwell's demons that can separate chemicals with little (although never zero) work. On page 1034 of this issue, Holt and co-workers (*1*) describe experiments that are a fascinating step toward the development of highly efficient membranes.

The separation of gases and liquids by membranes can often be more cost and energy effective than traditional separation methods, such as distillation or absorption. An ideal membrane would have excellent stability under a wide range of process conditions, high selectivity for the chemicals of interest, and also produce a large molecular flux with a small driving force. Most membranes in use industrially are polymeric, and the fabrication of these devices is highly developed. Polymeric membranes for gas separations show a near-universal tradeoff between flux and selectivity (*2*); materials with high throughput also have low selectivity, and vice versa. Polymeric membranes are also typically unsuitable for very high–temperature applications. New membrane materials that can overcome these fundamental hurdles could ultimately drastically reduce the energy consumed in present-day separation operations.

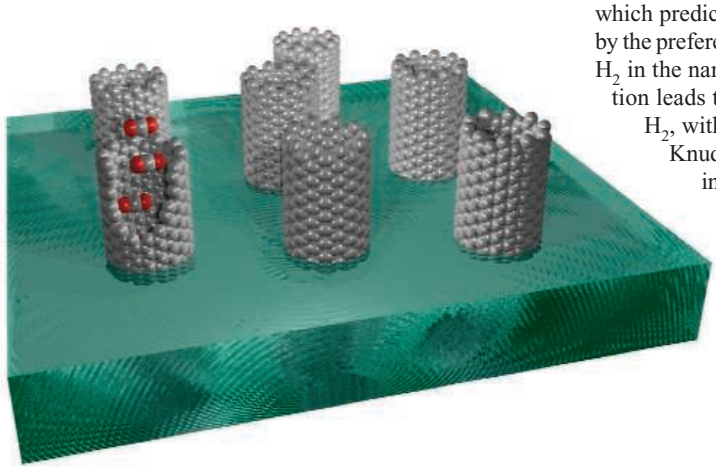
The possibility of using carbon nanotubes as membranes for gas separation has been recognized for some time. The first examinations of this idea used molecular dynamics (MD) simulations of gas transport inside single-walled nanotubes (*3*, *4*). These simulations predicted that the transport of gases inside nanotubes is orders of magnitude faster than in any other known materials with nanometer-scale pores. These rapid transport rates exist because the walls of nanotubes are much smoother (on atomic scales) than other materials.

These model predictions have now been tested experimentally by Holt *et al*., who have fabricated the first membranes from aligned single- and double-walled nanotubes. This work follows similar experiments by Hinds and co-workers with membranes made from multiwalled nanotubes (*5*, *6*). Both groups fabricated membranes in multistep processes, with the end result being nanotubes that penetrate a thin impermeable film (see the figure); Holt *et al*. used a silicon nitride matrix whereas Hinds *et al*. used a polymer. The pore diameters of nanotubes in the membranes made by Holt *et al*. are 1.3 to 2 nm, whereas the nanotubes used by Hinds *et al*. are considerably wider. Experimental observations of single-component gases permeating through both types of membranes show rapid transport of gases. The transport rates observed by Holt *et al*. for a range of gases are one to two orders of magnitude larger than would be predicted by assuming a Knudsen description, which is in quantitative agreement with predictions from simulations (*7*).

These experiments show that carbon nanotube membranes can have spectacularly high fluxes, but to be useful as membranes, they must also show high selectivity. The experiments to date have only examined single-component transport, so no direct information on this crucial issue is available. As is commonly the case for nanoporous membranes, estimation of membrane selectivities from single-component experiments is not expected to be accurate. Mixture selectivity has been studied with MD simulations of single-walled nanotube membranes for CH_4/H_2 mixtures (*8*), which predict that the selectivity is dominated by the preferential adsorption of CH_4 relative to H_2 in the nanotubes. This preferential adsorption leads to selective transport of CH_4 over H_2, with selectivities as high as 10 to 20. Knudsen transport of this gas mixture, in contrast, would give a selectivity of 2.8, favoring H_2 transport over CH_4.

The discussion above has focused on gas separations, but the need for efficient liquid separations is just as great. Both Holt *et al*. and Hinds and co-workers have performed experiments assessing water transport through their nanotube membranes. Similar to what is seen with gases, water is observed to move through the membranes extremely rapidly. The transport rates reported by Holt *et al*. are in good agreement with predictions made from MD simulations (*9*). Experimental studies of the selectivity of these membranes when they are applied to liquid mixtures will be of great interest.

The path to move from the elegant membranes fabricated by Holt *et al*. and Hinds *et al*. to devices suitable for large-scale applications—for example, CO_2 capture—will not be an easy one. The key challenge in this context is to scale up the fabrication techniques that have now been successfully demonstrated to economically produce membranes with large surface areas. The scope of this challenge is large, but the potential payoff is commensurately large. These experiments should also motivate other approaches to membrane fabrication that use carbon nanotubes as one component in a composite membrane rather than as the sole avenue for molecular transport across a membrane. So-called mixed matrix membranes that embed small nonpolymeric parti-



Faster flow. A schematic illustration of a membrane in which single-walled carbon nanotubes (gray) create avenues for transport of molecules (shown as red and gray balls passing through a nanotube) across an otherwise impermeable film (green).

CREDIT: A. SKOULIDAS/CARNEGIE MELLON UNIVERSITY

D. S. Sholl is in the Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA. E-mail: sholl@andrew.cmu.edu J. K. Johnson is in the Department of Chemical and Petroleum Engineering, University of Pittsburgh, Pittsburgh, PA 15260, USA. E-mail: karlj@pitt.edu

Published by AAAS

cles inside a polymeric matrix are a well-known route to improving the properties of polymeric membranes (*10*). Mixed matrix membranes hold a great economic advantage over "pure" inorganic membranes, because the inclusion of the inorganic component is, in principle, a relatively simple addition to existing methods for making large–surface area polymeric membranes. The experiments of Holt *et al*. strongly suggest that mixed matrix nanotube/polymer membranes may bring us closer to mass production of devices that do the job assigned to Maxwell's fictional demon.

References

1. J. K. Holt *et al*., *Science* **312**, 1034 (2006).
2. B. Freeman, *Macromolecules* **32**, 375 (1999).
3. A. I. Skoulidas, D. M. Ackerman, J. K. Johnson, D. S. Sholl, *Phys. Rev. Lett.* **89**, 185901 (2002).
4. V. P. Sokhan, D. Nicholson, N. Quirke, *J. Chem. Phys.* **117**, 8531 (2002).
5. B. J. Hinds *et al*., *Science* **303**, 62 (2003).
6. M. Majumder, N. Chopra, R. Andrews, B. J. Hinds, *Nature* **438**, 44 (2005).
7. A. I. Skoulidas, D. S. Sholl, J. K. Johnson, *J. Chem. Phys.* **124**, 054708 (2006).
8. H. Chen, D. S. Sholl, *J. Membrane Sci.* **269**, 152 (2006).
9. A. Kalra, S. Garde, G. Hummer, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 10175 (2003).
10. T. C. Merkel *et al*., *Science* **296**, 519 (2002).

10.1126/science.1127261

CHEMISTRY

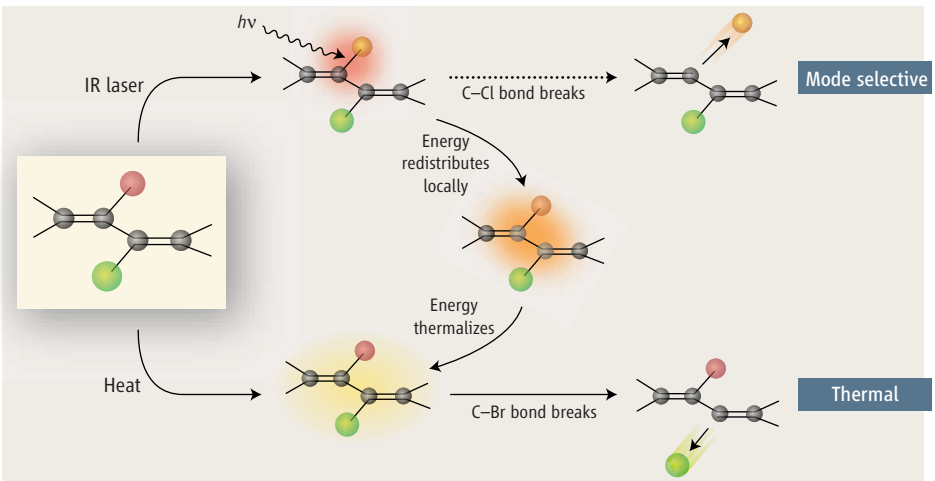
Mode-Selective Control of Surface Reactions

Researchers have achieved the goal of controlling chemical reactions by selectively exciting a single vibrational mode. A free-electron laser selectively desorbs H_2 from a silicon surface coated with hydrogen and deuterium.

John C. Tully

Mode-selective chemistry—the ability to energize a specific chemical bond and thereby promote a desired reaction pathway—has been a widely pursued goal. The vision of using a tunable infrared (IR) laser to divert a reaction from its dominant thermal pathway toward a desired product has even attracted commercial interest. For the most part, however, molecules have not cooperated. On page 1024 of this issue, Liu *et al*. (*1*) report convincing evidence for the mode-selective desorption of H_2 from a hydrogen-covered silicon surface.

The main impediment to IR mode-selective chemistry is that vibrational energy tends to be redistributed rapidly within a molecule. An initially excited, high-frequency localized mode can quickly de-excite by transferring its energy into combinations of lower frequency modes. In small molecules with sparse vibrational modes, few (if any) combinations of low-frequency modes can accept the energy, and the lifetime of the initially excited mode may be sufficiently long to allow mode-selective chemistry. Indeed, the outcome of the gas-phase reaction of H atoms with singly deuterated water (HOD) can be controlled through laser excitation of specific HOD vibrational modes (*2*, *3*). However, in large molecules, in condensed phases, and at surfaces, huge numbers of low-frequency modes can accept energy, and energy randomization is very rapid (generally on the picosecond time scale or faster). Thus, energy does not remain localized in a bond for a sufficiently long time to influence a chemical reaction. The resulting chemistry is thermal rather than selective: the weakest bond breaks or the most reactive site reacts (see the figure).



Conventional wisdom. The hypothetical molecular fragment shown here contains a weak bond between bromine (green) and carbon (gray) and a stronger bond between chlorine (red) and carbon. Heating results in thermal cleavage of the weakest bond, C–Br. IR excitation of the C–Cl vibration aims to selectively break this stronger bond. However, rapid intramolecular redistribution of the laser energy is likely, producing the same result as heating: the nonselective breaking of the weaker bond. The experiments of Liu *et al.* are more complex than this illustration, but nevertheless demonstrate a mode-selective, nonthermal pathway.

Attempts at IR mode-selective desorption of molecules from surfaces have been reported previously, but none have convincingly shown selectivity. A beautiful series of experiments was carried out by T. J. Chuang and co-workers in the early 1980s. In one example (*4*), they observed greatly enhanced desorption yields of NH_3 from copper and silver surfaces when the IR laser was tuned to resonance with an ammonia vibrational mode. However, when they co-adsorbed NH_3 and ND_3 on the surface and tuned the laser to an NH_3 vibrational frequency, both NH_3 and ND_3 desorbed in statistical (nonselective) proportions; NH_3 did not desorb preferentially. Thus, energy is deposited only when the laser frequency is resonant with a vibration, but the energy rapidly randomizes, and the subsequent chemistry is driven by nonselective heating. Redlich *et al*. recently reported similar findings for isotopic mixtures of methane physisorbed on an NaCl surface; again, isotopes were found to desorb in statistical proportions irrespective of which isotope was vibrationally excited (*5*).

The experimental findings of Liu *et al*. (*1*) are in striking contrast with the wavelength-dependent but nonselective desorption found in these previous studies. The authors first created an adsorbed layer of about 15% H atoms and 85% D atoms on an Si(111) surface. They then irradiated the surface with a free-electron laser tuned to the 4.8-μm Si-H stretching mode. They found that almost all desorbing atoms were H_2; less than 5% of desorbing molecules were HD or D_2. This startling result rules out any local heating mechanism, which would produce a statistical mixture (2% H_2, 26% HD, and 72% D_2). Liu *et al*. have thus clearly

The author is in the Departments of Chemistry, Physics, and Applied Physics, Yale University, New Haven, CT, 06520, USA. E-mail: john.tully@yale.edu

achieved IR mode-selective chemistry. The atomic-level mechanism underlying this process is far from clear, however.

The authors observe a quadratic dependence of desorption yield on laser intensity. It is tempting to implicate two neighboring H-Si units, each vibrationally excited by a single IR photon. Liu *et al*. correctly caution, however, that a quadratic dependence on intensity does not necessarily imply a two-photon process. The measured activation energy for thermal desorption of H_2 from H-Si(111) is 2.4 eV (*6*)—far more than two 0.26-eV IR photons. The measured activation energy of a chemical reaction does not necessarily equal the height of an actual energy barrier. Nevertheless, it is difficult to see how only two photons can produce desorption. Any atomic trajectory that might be launched by IR laser excitation of two neighboring-singly excited H-Si bonds can also be produced by heating—there cannot be any secret passages that are accessed only by IR laser excitation. Therefore, if only two photons worth of energy can produce desorption in the laser experiment, then the same amount of energy can produce desorption thermally. But if so, the activation energy would be far lower than 2.4 eV.

It thus appears necessary to invoke many IR photons. Perhaps there is some contribution from vibrational energy pooling. Chang and Ewing have observed this effect in physisorbed layers of CO on NaCl (*7*), where dipole-dipole coupling induces many molecules excited by a single photon to pool their energy into a few highly vibrationally excited molecules (*8*). For the H-Si(111) system, however, this effect is expected to be much less important; the dipole-dipole coupling is weaker than for CO, and the vibrational lifetime is much shorter.

It appears more likely that the high intensities achievable with the free-electron laser produce multiple (sequential) photon absorption by individual H-Si bonds. Much of this energy must remain localized long enough for two H atoms to combine and surmount the 2.4-eV barrier. Scanning tunneling microscopy studies of the H-Si(100) surface have shown that energetic electrons are far more likely to induce the breaking of an Si-H bond than the breaking of an Si-D bond (*9*). This difference is believed to be the reason for the increased lifetime of semiconductor devices upon deuterium substitution (*10*). Van de Walle and Jackson (*11*) have proposed that energetic electrons excite both Si-H and Si-D vibrations. However, the Si-D vibrational frequencies more closely match those of the silicon substrate, and energy dissipation should thus occur more rapidly for Si-D, leading to preferential breaking of Si-H bonds.

The experiments of Liu *et al*. may exhibit similar behavior, where energy transferred to Si-D units quickly dissipates, leaving only the Si-H units energized. But whatever the mechanism, Liu *et al*. have successfully accomplished a long-standing goal: IR mode-selective chemistry in a many-atom system.

References

1. Z. Liu *et al.*, *Science* **312**, 1024 (2006).
2. A. Sinha, M. C. Hsiao, F. F. Crim, *J. Chem. Phys.* **92**, 6333 (1990).
3. M. J. Bronikowski, W. R. Simpson, R. N. Zare, *J. Phys. Chem.* **97**, 2194 (1992).
4. I. Hussla *et al.*, *Phys. Rev.* B **32**, 3489 (1985).
5. B. Redlich *et al.*, *J. Chem. Phys.* **124**, 044704 (2006).
6. G. A. Reider, U. Höfer, T. F. Heinz, *J. Chem. Phys.* **94**, 4080 (1991).
7. H.-C. Chang, G. E. Ewing, *Phys. Rev. Lett.* **65**, 2125 (1990).
8. S. A. Corcelli, J. C. Tully, *J. Phys. Chem.* A **106**, 10849 (2002).
9. Ph. Avouris *et al.*, *Surf. Sci.* **363**, 368 (1996).
10. J. W. Lyding, K. Hess, I. C. Kizilyalli, *Appl. Phys. Lett.* **68**, 2526 (1996).
11. C. G. Van de Walle, W. B. Jackson, *Appl. Phys. Lett.* **69**, 2441 (1996).

10.1126/science.1126341

APPLIED PHYSICS

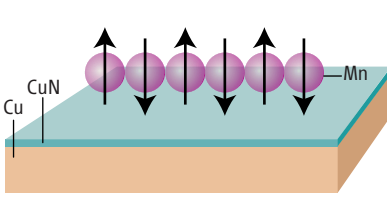
Assembly and Probing of Spin Chains of Finite Size

Understanding magnetic ordering at the atomic scale is essential for spintronic technology. A linear chain of manganese atoms has been created for studying one-dimensional systems.

Harald Brune

The creation, investigation, and manipulation of low-dimensional model systems is of fundamental importance in condensed-matter physics. Moreover, an understanding of the wide variety of electronic and magnetic properties of these models—and their associated phase transitions—may lead to applications in spintronics and other areas of device physics. To an increasing degree, such model systems have been created by researchers in surface science. A well-known example was achieved in 1993 by Eigler and co-workers at IBM Almaden, who used a low-temperature scanning tunneling microscope (STM) to arrange adsorbed atoms into a corral that imposed a circular boundary on the surface-state electrons of the underlying single crystal (*1*). The resulting quantum interference patterns exactly displayed the solution of the Schrödinger equation, which for that geometry can be given in analytical form. A second but no less exciting example is now reported from the same lab on page 1021 of this issue by Hirjibehedin *et al*., who have carried out low-temperature STM measurements of manganese atom chains (of up to 10 atoms), assembled by atomic manipulation on copper nitride islands that provide an insulating monolayer between the chains and a copper substrate (*2*). These chains are model systems representing one-dimensional (1D) Heisenberg antiferromagnets of finite size.

A Heisenberg chain is a linear arrangement of spins S mutually coupled by an exchange interaction with energy J. Several order-disorder transitions are predicted for such chains. In the absence of magnetic anisotropy, for example, long-range order is predicted to be lost at any finite temperature (*3*, *4*). A model system for a ferromagnetic Heisenberg chain (where all the spins are aligned parallel to each other) has been realized in the form of atomic cobalt chains created by step decoration of vicinal platinum single-crystal surfaces. In this case, anisotropy was found to stabilize small ferromagnetic spin blocks (*5*). 1D ferrimagnets (where neighboring spins are antiparallel but do not cancel, leaving a net moment) have been realized with molecular magnets (*6*) and have been found to display the predicted slow magnetization relaxation (*7*). However, a chain with antiferromagnetic coupling (where neighboring spins are antiparallel) is of fundamental importance in many-body physics, as it is one of the few systems where a nontrivial many-particle ground state is known exactly (*8*). The quantum mechanical nature of the spins gives rise to the collapse of the Néel state (the arrangement of antiparallel spins) into a single



Model magnet. Schematic representation of an antiferromagnetic chain of six manganese atoms residing on a monolayer of insulating copper nitride grown on a Cu(100) substrate.

The author is at the Institut de Physique des Nanostructures, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland. E-mail: harald.brune@epfl.ch

wave function of the entire chain, which is subject to quantum fluctuations and displays quantum phase transitions in high magnetic fields (*9*).

Until now, model systems of 1D Heisenberg antiferromagnets have been bulk crystals with electronically coupled quasi-1D chains, such as $CuGeO_3$, copper tetraminsulfate, $CsNiCl_3$, or copper benzoate $Cu(C_6D_5COO)_2{\cdot}3D_2O$ (*10*), which have been investigated by electron spin resonance, nuclear magnetic resonance, magnetic susceptibility, and inelastic neutron scattering. Similar to the quasi-1D bulk samples, the manganese atom chains are strongly coupled along the chain yet weakly coupled to the environment. In contrast to the bulk samples, the chains have finite length, but structures with arbitrary geometries can be assembled, and the nearest neighbor distance as well as the adsorption site can be freely chosen. The STM allows direct measurement of the chain properties for each configuration—for instance, by inducing spin-flips and total spin changes that appear as steps in differential conductance curves (dI/dV, where I is current and V is voltage).

Odd chains have a zero-bias dip in dI/dV caused by spin-flip excitation (change in magnetic quantum number m); all chains have conductance steps arising from the total spin change. The size of the conductance steps in the even chains can reach up to one order of magnitude. The absence of spin-flip excitations for even chains shows that the total spin in the ground state is $S_{tot} = 0$, and the presence of such excitations in odd chains implies their ground state to have $S_{tot} \neq 0$; therefore, the chains are ordered antiferromagnetically. The interatomic coupling strength J of the chains is measured by means of the energy difference between ground and first excited state of a dimer at zero field. The influence of the chemical environment on J is clearly evident, as Hirjibehedin *et al.* find J = 6.2 meV in chains placed on Cu atoms of the insulating CuN layer, but J = 3.0 meV for a chain placed on nitrogen atoms. In the presence of a magnetic field, the total spin transition of a dimer splits up into three energies, corresponding to a transition from $S_{tot} = 0$ to $S_{tot} = 1$ with quantum numbers $m = 0, \pm 1$. From the energy of the total spin change of a linear trimer, the authors deduce $S_{tot} = S = \frac{1}{2}$. Because J and S are known, the energies for the total spin transitions for all chain lengths n can be predicted with a Heisenberg open-chain model. The peak positions up to n = 6 of this model are in excellent agreement with experiment. Moreover, the authors find inelastic electron tunneling spectroscopy (IETS) selection rules for spin transitions in the chains: $\Delta S_{tot} = 0, \pm 1$, and $\Delta m = 0, \pm 1$, but not ± 2.

These results represent an important step in the creation, understanding, and manipulation of low-dimensional spin systems. Understanding the selection rules is a theoretical challenge, and the role of magnetic anisotropy needs to be highlighted. Further experimental progress should be possible with spin-polarized STM, which might reveal whether internal order is indeed absent or may allow direct excitation of magnetic transitions with the injection of spin-polarized currents. The most exciting part of this report is that one can now create and explore any arrangement of spins, in particular those less likely to be found in nature's crystals.

References and Notes

1. M. F. Crommie, C. P. Lutz, D. M. Eigler, *Science* **262**, 218 (1993).
2. C. F. Hirjibehedin, C. P. Lutz, A. J. Heinrich, *Science* **312**, 1021 (2006); published online 30 March 2006 (10.1126/science.1125398).
3. N. D. Mermin, H. Wagner, *Phys. Rev. Lett*. **17**, 1133 (1966).
4. P. Bruno, *Phys. Rev. Lett*. **87**, 137203 (2001).
5. P. Gambardella *et al*., *Nature* **416**, 301 (2002).
6. A. Caneschi *et al*., *Angew. Chem. Int. Ed*. **40**, 1760 (2001).
7. R. J. Glauber, *J. Math. Phys*. **4**, 294 (1963).
8. H. Bethe, *Z. Phys*. **71**, 205 (1931).
9. M. Enderle *et al*., *Physica B* **276**, 560 (2000).
10. D. C. Dender *et al*., *Phys. Rev. B* **53**, 2583 (1996).

11. The author gratefully acknowledges discussions with H. Rønnow.

10.1126/science.1127387

BEHAVIOR

Foresight and Evolution of the Human Mind

Planning for the future is a fundamental human survival strategy. New results suggest that great apes can anticipate future needs and that this ability has roots more ancient than previously thought.

Thomas Suddendorf

What is in your pockets? Chances are you carry keys, money, cosmetics, a Swiss Army knife, or other tools—because they may be useful at some future point. Humans have the ubiquitous capacity to imagine, plan for, and shape the future (even if we do frequently get it wrong). This capacity must have long been of major importance to our survival (see the figure) and may have been a prime mover in human cognitive evolution (*1*). Stone toolkits and spears from archaeological finds suggest that the ancestors of modern humans already prepared for the future hundreds of thousands of years ago. On page 1038 of this issue, Mulcahy and Call (*2*) show that the roots of these abilities may go much deeper still. Though great apes have not invented containers to carry tools, the experiments demonstrate that they can save tools for future use.

Enhanced online at www.sciencemag.org/cgi/content/full/312/5776/1006

Of course, other animals also act in ways that increase their chances of future survival. Many species have evolved preparatory instincts that lead them, for example, to build nests or hoard food. Associative learning mechanisms further allow individuals, rather than entire species, to predict recurrences on the basis of cues (for example, a smell signaling food). But animals are not mere associative automatons. Recent evidence shows that some can make causal inferences (*3*, *4*). Great apes even seem capable of imagining situations they cannot directly perceive (*5*). They can also make simple tools to solve nearby problems, such as fashioning an appropriate stick to obtain food that would otherwise be out of reach (*6*). Yet there seems little evidence that animals ponder the more distant future. Wolfgang Köhler, who first documented "insightful" problem-solving in chimpanzees nearly 90 years ago, concluded that "The time in which the chimpanzee lives is limited in past and future" (*6*). Recent reviews concur that animal foresight more generally may be profoundly restricted (*7*, *8*), even though innovative work on scrub-jays suggests that these birds have some surprising competence in the specific context of food caching (*9*). In this week's *Science* Express, Dally *et al.* (*10*), for example, report that scrub-jays adjust their caching behavior in ways that effectively decrease the chances that the food will be stolen by others. They appear to keep track of what caching was observed by which other jay and move food to new locations accordingly. However, the caching behaviors may be driven by a combination of predispositions and specific learning algorithms rather than by reasoning about past and future (*10*, *11*). It remains unclear what exactly the limits are for animal foresight.

The author is in the School of Psychology, Early Cognitive Development Unit, University of Queensland, Brisbane, QLD, 4072 Australia. E-mail: t.suddendorf@psy.uq.edu.au

Tool carriers. (**Left**) Humans routinely prepare for the future, as illustrated here by the many tools the "ice-man" Ötzi carried through the Alps more than 5000 years ago. (**Right**) Some great apes also use tools and, in the laboratory at least, can transport them for future use.

Mulcahy and Call's experiments demonstrate the most extensive foresight yet in nonhuman primates. Bonobos and orangutans at the Wolfgang Köhler Primate Center in Leipzig, Germany, first learned to use an object as a tool to obtain grapes from an apparatus in a test room. They were then presented with objects that would make suitable and unsuitable tools for this task, from which they could select any to take into a waiting room (where they could still see the apparatus). An hour later, they were allowed back into the test room. In almost half of the trials, subjects spontaneously selected a suitable tool, transported it out to the waiting room, and then back into the test room to get the reward. They also returned with unsuitable tools, but significantly less so than with suitable ones. In a second experiment, the best performing bonobo and orangutan were given the tool choice in the evening but were only allowed to return to the test room the next morning. Impressively, both still returned with suitable tools in more than half of the trials. In a third experiment, apes had to first learn to use a hook to obtain grape juice from a different apparatus. This apparatus was then removed before subjects were given the opportunity to select tools, and then reinstalled only after they had returned from the waiting room. The apes still performed reasonably well even though they had to make their choices without being able to see the reward apparatus.

In an important final control condition, the apparatus was not reinstalled at all. Subjects were still given the reward if they returned with a suitable tool, but they could not use the tool on the apparatus. If the apes simply associated the tool with the reward, performance in this condition should be equivalent to that in the third experiment. Subjects returned less frequently with the suitable tool, however, leading the authors to conclude that performance in the other experiments reflected planning rather than simple associative learning.

The star performer was the orangutan Dokana, who in the first study succeeded 15 out of 16 times—more than twice as often as the second best performing animal. Dokana also did well in the second and third experiments. Unfortunately, Dokana did not participate in the final control experiment. Instead, it was administered to a naïve group of four apes, raising concern that this might account for poorer performance in the control condition. As the authors acknowledge, two of the control animals never returned to the test room with the "suitable" tool and, hence, had no opportunity to learn from the reward. The other two performed as well as two of the six apes tested in the first experiment, returning in two and five trials with suitable tools. Clearly, more research is needed to entirely rule out associative explanations. Still, the findings suggest that "the time in which Dokana lives spans more than a night."

Köhler saw no sign that chimpanzees consider the future beyond striving to satisfy current needs. This led to the proposal that animal foresight may be limited by an inability to entertain future need or drive states that they do not currently experience—such as imagining being thirsty when one is quenched (*1*, *7, 12*). There would be little point considering the remote future if one is driven merely by a desire to satisfy current needs. To examine this hypothesis, two-room tests have been proposed in which human or animal subjects can prepare to secure a need that they do not currently experience (*11*, *13*). Mulcahy and Call's study is similar in structure, but did not measure or control subjects' motivational states. It is probably fair to assume that the apes desired grape rewards throughout the duration of the experiment. Thus, although the data suggest anticipation of the future need for a tool, they do not necessarily imply anticipation of a future state of mind.

Nonetheless, Mulcahy and Call's results are ground-breaking and may even prove to be as seminal as Premack and Woodruff's first evidence that chimpanzees may have a theory of mind (*14*). But unequivocally establishing complex mental capacities in nonverbal animals is difficult, as decades of subsequent theory of mind research testify. Much clever experimentation is required to determine what foresight apes have and what the limits of this ability are.

The stakes are high. By identifying what capacities our closest living relatives share with us, we can get a glimpse at our evolutionary past. It is more parsimonious to conclude, as the authors do, that traits shared with all great apes are inherited from our common ancestor, rather than that they evolved independently. Even without fossils, the mental capacities of this ancestral species are becoming increasingly clear (*5*). This provides a starting point from which we can begin to reconstruct the evolution of the human mind.

There is still much to learn about human foresight. Researchers have concentrated more on memory than on anticipation. Yet, it is accurate prediction of the future, more so than accurate memory of the past per se, that conveys adaptive advantages. Emerging research suggests that remembering past episodes and imaging future events may be two sides of the same faculty [so-called mental time travel (*1*)]; various links between the two have been reported. Amnesic patients who are unable to answer simple questions about yesterday's events have been found to be equally unable to say what might happen tomorrow (*15*), and children begin to accurately answer both such questions from around the same age (*11*). Imagining future events and remembering past events are associated with similar brain activity (*16*), and factors like temporal distance influence in similar ways the reported phenomenological characteristics of both (*17*).

But many questions remain. What are the mechanisms of future-oriented cognition? How do they develop? How can performance be disturbed or enhanced? It is time to carefully investigate foresight, because humans may need to get better at it if we are to continue to survive.

References

1. T. Suddendorf, M. C. Corballis, *Genet. Soc. Gen. Psychol. Monogr.* **123**, 133 (1997).
2. N. J. Mulcahy, J. Call, *Science,* **312**, 1038 (2006).
3. J. Call, *J. Comp. Psychol.* **118**, 232 (2004).
4. A. P. Blaisdell, K. Sawa, K. J. Leising, M. R. Waldmann, *Science* **311**, 1020 (2006).
5. T. Suddendorf, A. Whiten, *Psychol. Bull.* **127**, 629 (2001).
6. W. Köhler, *The Mentality of Apes* (Routledge & Kegan Paul [(Original work published 1917), London, 1927)].
7. W. A. Roberts, *Psychol. Bull.* **128**, 473 (2002).
8. T. Suddendorf, J. Busby, *Trends Cognit. Sci.* **7**, 391 (2003).
9. N. S. Clayton, T. J. Bussey, A. Dickinson, *Nat. Rev. Neurosci.* **4**, 685 (2003).
10. J. M. Dally, N. J. Emery, N. S. Clayton, *Science* 18 May 2006 (10.1126/science.1126539).
11. T. Suddendorf, J. Busby, *Learn. Motiv.* **36**, 110 (2005).
12. N. Bischof, *Das Rätzel Ödipus* [*The Oedipus Riddle*] (Piper, Munich, 1985).
13. E. Tulving, in *The Missing Link in Cognition*, H. S. Terrace, J. Metcalfe, Eds. (Oxford Univ. Press, Oxford, 2005), pp. 3–56.
14. D. Premack, G. Woodruff, *Behav. Brain Sci.* **1**, 515 (1978).
15. S. B. Klein, J. Loftus, J. F. Kihlstrom, *Soc. Cognition* **20**, 353 (2002).
16. J. Okuda *et al.*, *Neurolmage* **19**, 1369 (2003).
17. A. D'Argembeau, M. Van der Linden, *Consciousness and Cognition* **13**, 844 (2004).

10.1126/science.1129217

NEUROSCIENCE

Gatekeeper at the Synapse

Motor neurons innervate muscles through structures containing calcium channels and vesicles poised to release transmitter. A newly described protein organizes these structures.

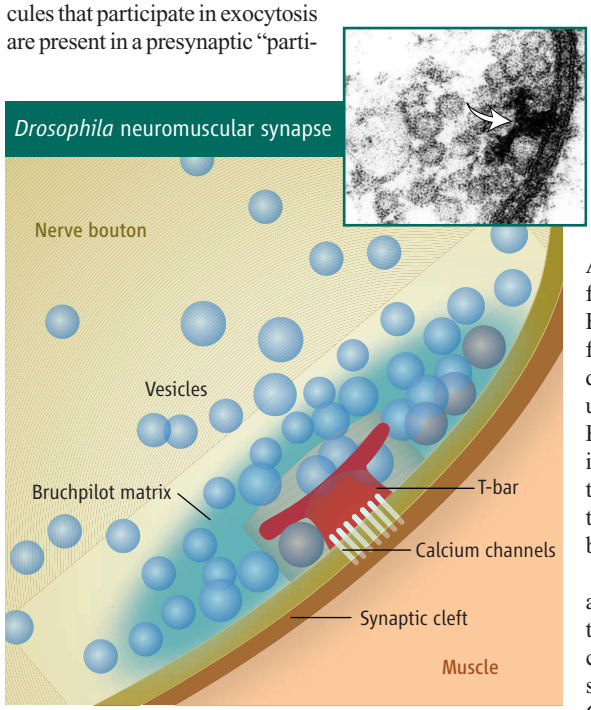
Harold L. Atwood

The most important mechanism for rapid cell-to-cell communication within the nervous system is the synapse, where neuron meets target for the relay of chemical messages. Fast neurotransmission not only requires very close apposition of presynaptic and postsynaptic partners, but also necessitates a precise structural arrangement of cellular components on both sides of the synaptic cleft to facilitate effective signal transmission. On page 1051 of this issue, Kittel *et al.* (*1*) report how a recently discovered presynaptic protein acts as a gatekeeper for those components that need to be assembled for fast release of neurotransmitter molecules from synaptic vesicles.

Although synapses of different neurons are tuned to meet the demands of specific activities and are correspondingly diverse in structure and function (*2*), they share several basic features. For fast neurotransmission, apposing presynaptic and postsynaptic membranes are typically separated by 10 to 50 nm, rigidly aligned, and held in place by adhesion molecules. Postsynaptic neurotransmitter receptors are thus rapidly activated with very little time lost in the diffusion of signals across the synaptic cleft (*3*). All fast chemical synapses possess voltage-gated calcium channels in the presynaptic membrane such that calcium ions entering the presynaptic nerve terminal trigger rapid release, or exocytosis, of neurotransmitter from synaptic vesicles (*4*). The needs of rapid communication are met by very close association between calcium channels and presynaptic vesicles. The time between the entry of calcium ions into the presynaptic terminal and the onset of the postsynaptic potential is about 200 μs at the squid giant synapse (*5*). For release of transmitter to occur in this brief time, the synaptic vesicle must be "docked" at the presynaptic membrane and close to one or more calcium channels. These channels, when opened by a presynaptic nerve impulse, create a small domain of relatively high calcium concentration, which activates exocytosis (*4*, *6*).

Precise positioning of presynaptic calcium channels and docked vesicles is ensured by cytoskeletal filaments linking these participants. At the frog neuromuscular junction, electron microscope tomography and scanning electron microscopy have revealed a regular cytoskeletal lattice tethering a linear array of synaptic vesicles to large transmembrane particles, the putative calcium channels (*7*). The lattice is part of an electron-dense structure seen at the presynaptic membrane by standard transmission electron microscopy. Collectively, the calcium channels, docked vesicles, and tethering elements are known as the presynaptic active zone, because this limited region is the site of impulse-evoked exocytosis at the synapse (*8*, *9*).

At mammalian synapses of the central nervous system, the active zone is organized as a presynaptic grid (*10*) rather than a linear array. Pre- and postsynaptic membranes are tightly linked by adhesion molecules, whereas molecules that participate in exocytosis are present in a presynaptic "particle web" (*11*). Yet more elaborate active-zone structures occur at vertebrate sensory synapses as very prominent presynaptic ribbons, which are thought to convey vesicles continuously to the sites of exocytosis (*12*). The molecular components of these specialized structures have recently been described (*13*). Presynaptic densities differ in form among diverse synapses, but they share functionally important features that are being closely studied for clues to the regulation of synaptic performance.

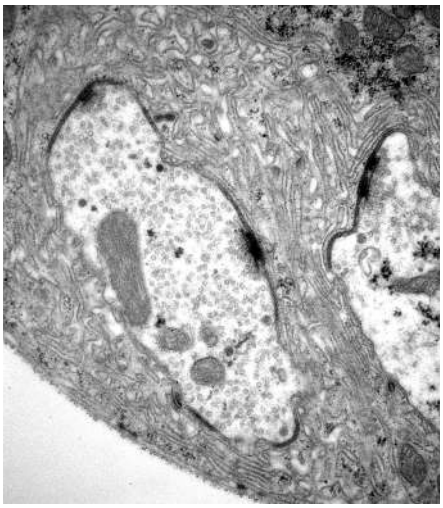


The Bruchpilot matrix. Transmission electron micrograph of a presynaptic density (**top**, **right**) illustrates assembly of synaptic vesicles at the presynaptic density, or T-bar, in a *Drosophila* synapse. The synaptic vesicles are about 40 nm in diameter. The diagram illustrates the *Drosophila* T-bar surrounded by a donut-shaped zone of Bruchpilot protein, its cluster of calcium channels, and several vesicles attracted to the active zone. The thickness, substructure, and lateral extent of the Bruchpilot matrix are not yet fully defined.

Emerging evidence for the identity of structural proteins at vertebrate sensory synapses raises the question of their occurrence in other synapses. The structural proteins Bassoon and CAST (cytomatrix of the active zone–associated structural protein) are closely associated with the presynaptic active zone of mammalian hippocampal synapses (*14*). But counterparts for these proteins have not been found in other groups of organisms, including arthropods. Kittel and colleagues have now identified a candidate for a missing functional homolog. Active zones with presynaptic densities at invertebrate synapses are of particular interest because of the emergence of the *Drosophila* larval neuromuscular junction as a system in which modern genetic techniques can be readily applied to elucidate functional properties of specific molecules. An organizing role has been defined for a specific molecule named Bruchpilot (German for "crash pilot"; flies in which the protein has been depleted by RNA interference are unable to sustain proper flight) (*15*). Kittel *et al.* show that this organizing molecule is required for formation and anchoring of the presynaptic density, but it does not appear to be an integral component.

Bruchpilot is a structural protein apparently present at all synapses of the *Drosophila* nervous system. It contains coiled-coil domains and shares homology with vertebrate CAST, a protein that interacts in a complex assemblage with other vertebrate presynaptic proteins (including Bassoon, Munc 13-1, and RIM1) all closely associated with the active zone. At neuromuscular junctions of *Drosophila* mutants lacking Bruchpilot, the presynaptic densities (so-called "T-bars") disappear. Despite this, nerve-evoked transmission still occurs but is reduced to about 25% of normal, and the timing of exocytosis is less precise. Spontaneous release of the transmitter is not affected, so the fundamental process of exocytosis is preserved. Presynaptic calcium channels are reduced in number and are less tightly clustered and less closely associated with docked vesicles. Apparently, Bruchpilot is not part of presynaptic density itself but probably surrounds it as a matrix (see the first

The author is in the Department of Physiology, University of Toronto, Toronto, Ontario M5S 1A8, Canada. E-mail: h.atwood@utoronto.ca



***Drosophila* neuromuscular junction.** Transmssion electron micrograph of two presynaptic boutons, wrapped by processes of a muscle cell. Each has two prominent T-bars and associated clusters of synaptic vesicles. The bouton at the left is approximately 1 μm wide and 2.5 μm long.

CREDIT: L. MARIN/UNIVERSITY OF TORONTO

figure), and is required (directly or indirectly) for normal localization of docked vesicles and calcium channels. It probably acts as a constraining guardian and gatekeeper for the active zone. At present, the form of the matrix is not well defined; electron micrographs show that it does not prevent the movement of synaptic vesicles into the zone near the presynaptic membrane. Thus, the matrix may consist of fine filaments of Bruchpilot protein on the cytoplasmic surface of the presynaptic membrane.

Arrangement of the Bruchpilot matrix at the active zone is revealed by a new technique, stimulated emission depletion (STED) microscopy (*16*), that shows fluorescence-labeled structures in biological specimens at higher resolution than with confocal microscopy. The Bruchpilot matrix is a doughnut-shaped structure with one or two prominent holes large enough to accommodate presynaptic densities and a cluster of calcium channels. Available structural evidence from transmission electron microscopy and freeze-fracture images (*17*, *18*) indicates that the *Drosophila* active zone normally consists of calcium channels and docked vesicles arranged along a T-bar (see the second figure). A ring of the Bruchpilot protein surrounds this aggregation, giving the doughnut-like appearance seen with STED. This structural arrangement suggests that the presynaptic density acts as an attractor of synaptic vesicles and a guide to their optimal disposition for fast exocytosis. It is also possible that the presynaptic density is required for clustering calcium channels at the active zone (see the figure). Without the Bruchpilot matrix, the presynaptic density does not form properly, and the effectiveness of nerve-evoked transmission is much reduced.

Whether the Bruchpilot matrix is directly responsible for clustering of calcium channels, or whether it is indirectly involved through attracting and anchoring the presynaptic density, remains an open question. Undoubtedly, the *Drosophila* T-bar and similar presynaptic active-zone densities of other arthropods have functions in common with vertebrate synaptic grids and ribbons (*11*, *12*), but they possess different components and anchoring molecules. Though destined never to follow a course of flight, the crash-pilot flies nonetheless have pointed to a new direction for analyzing synaptic structure and function.

References

1. R. J. Kittel *et al.*, *Science* **312**, 1051 (2006); published online 13 April 2006 (10.1126/science.1126308).
2. H. L. Atwood, S. Karunanithi, *Nat. Rev. Neurosci.* **3**, 497 (2002).
3. B. Katz, R. Miledi, *Proc. R. Soc. London Ser. B* **161**, 483 (1965).
4. E. F. Stanley, *Trends Neurosci.* **20**, 404 (1997).
5. R. Llinas, I. Z. Steinberg, K. Walton, *Proc. Natl. Acad. Sci. U.S.A.* **73**, 2918 (1976).
6. G. J. Augustine, E. M. Adler, M. P. Charlton, *Ann. N.Y. Acad. Sci.* **635**, 365 (1991).
7. M. L. Harlow, D. Ress, A. Stoschek, R. M. Marshall, U. J. McMahan, *Nature* **409**, 479 (2001).
8. J. E. Heuser, T. S. Reese, D. M. D. Landis, *J. Neurocytol.* **3**, 109 (1974).
9. R. Couteaux, M. Pécot-Dechavassine, *C. R. Acad. Sci. Ser. D* **271**, 2346 (1970).
10. K. Pfenninger, K. Akert, H. Moor, C. Sandri, *J. Neurocytol.* **1**, 129 (1972).
11. G. R. Phillips *et al.*, *Neuron* **32**, 63 (2001).
12. D. Lenzi, H. Von Gersdorff, *BioEssays* **23**, 831 (2001).
13. S. tom Dieck *et al.*, *J. Cell Biol.* **168**, 825 (2005).
14. A. Hagiwara, Y. Fukazawa, M. Deguchi-Tawarada, T. Ohtsuka, R. Shigemoto, *J. Comp. Neurol.* **489**, 195 (2005).
15. D. A. Wagh *et al.*, *Neuron* **49**, 833 (2006).
16. S. W. Hell, *Nat. Biotechnol.* **21**, 1347 (2003).
17. A. Prokop, I. A. Meinertzhagen, *Semin. Cell Dev. Biol.* **17**, 20 (2006).
18. C. J. Feeney, S. Karunanithi, J. Pearce, C. K. Govind, H. L. Atwood, *J. Comp. Neurol.* **402**, 197 (1998).

10.1126/science.1128445

CHEMISTRY

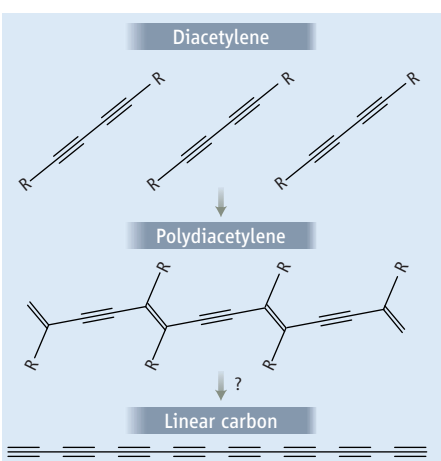
Dangerously Seeking Linear Carbon

Ray H. Baughman

It has been problematic to produce linear carbon molecules because key reactants are highly unstable and can explode. A solid state polymerization reaction avoids this problem and allows synthesis of these elusive products.

Three decades ago, there was an explosion in my laboratory at Allied Chemical. We were trying to make a new phase of carbon, called linear carbon, by thermally polymerizing diiododiacetylene (C_4I_2) crystals and then eliminating the iodine substituent. The first step yielded less than a gram of polymerized C_4I_2 crystals, which had a metallic luster. Intrigued by this sample, a talented chemist forgot caution for an instant and tapped the side of a sealed ampoule with his finger. The resulting explosion of the polymerized crystals caused minor injury and spelled the end of our efforts to make linear carbon.

Others have persisted, however, and thanks to their efforts a route to linear carbon is now coming closer to possible success. On page 1030 of this issue (*1*), Sun *et al.* report the synthesis of poly(C_4I_2) by a modified solid-state polymerization route. The authors use crystals in which C_4I_2 is complexed with a specially chosen agent, an oxalamide. In these crystals, 1,4-addition polymerization of diacetylene occurs spontaneously to produce single crystals of the complexed polydiacetylene (see the first figure). The complexing agent appears to stabilize the polymerized C_4I_2 against explosive decomposition (*2*) and eliminates the type of disorder that Sun *et al.* (*1*) report for crystals of pure C_4I_2. Their approach draws on a deep understanding of the application of cocrystallized molecules to pro-



Toward linear carbon. (**Top**) An array of diacetylene molecules in a polymerizable diacetylene single crystal. (**Middle**) A resulting polydiacetylene chain. (**Bottom**) A chain of linear carbon that might result from the extrusion of substituent groups R, such as iodine, from the polydiacetylene chain.

The author is at the Nanotech Institute, University of Texas at Dallas, Richardson, TX 75080, USA. E-mail: ray.baughman@ utdallas.edu

vide the diacetylene array geometry needed for polymerization (*3*, *4*).

The diacetylene solid-state polymerization reaction exploited by Sun *et al.* is a very special reaction, whose nature and importance were first established by the pioneering work of Gerhard Wegner (*5*). In some cases, this reaction proceeds completely topochemically; that is, the monomer structure accurately guides the continuous transformation from monomer to polymer. As a result, a monomer crystal of any size converts to a polymer crystal of similar size and structural perfection (see the second figure).

In the old efforts at Allied Chemical and elsewhere to commercially exploit polydiacetylene single crystals, many interesting properties were found (*6*, *7*), including some still being pursued for applications (*8*). The crystals have a per-chain stiffness close to that of diamond, very high coefficients for tripling the frequency of incoming light, and negative thermal expansion coefficients tunable to near zero by introducing defects. They have been deployed as mechanical stress sensors and as temperature-limit and chemical-exposure indicators. Also, the very high observed electron mobility for polydiacetylene single crystals suggests electronic device applications. Many of these properties deserve investigation for the polymer produced by Sun *et al*., but probably for fundamental rather than practical reasons, because poly(C_4I_2) chains lack the stability of more typically investigated polydiacetylenes with large organic substituents.

In perhaps the most important commercial application of diacetylene polymerization, printed inks containing diacetylene microcrystals have been used as time-temperature indicators. Diacetylenes typically have an actuation energy for thermal polymerization of about 20 to 28 kcal/mol, which can be tuned to closely match the activation energy and degradation rate for important perishables, from vaccines to foodstuffs. Color changes of the diacetylene microcrystals during thermal polymerization provide a visual indication of whether a perishable product in the same thermal environment has degraded as a result of integrated time-temperature exposure (*9*). Over a billion of these diacetylene polymerization–based indicators have been used on individual vaccine vials since 1996 to assist disease eradication in parts of the world that do not have a reliable cold chain, by helping to ensure the delivery of viable vaccines (*10*). The United States Army is using these diacetylene indicators on cartons of their MREs (Meals Ready to Eat).

Is it possible to make linear (two-coordinated) carbon as a distinct crystal phase that complements the known phases of four-coordinated carbon (cubic and hexagonal diamond) and of three-coordinated carbon (graphite, fullerenes, and carbon nanotubes)? Much of the literature for solid forms of linear carbon (also called carbyne or carbene) is difficult to interpret. Numerous reports claim the synthesis of diverse crystalline phases of linear carbon (*11*), but many are likely due to artifacts. In any case, no crystallographic structure determination has been reported for crystals containing only chains of linear carbon.

With increasing n, crystals comprising only polyynes—that is, R-$(C{\equiv}C)_n$-R molecules—become increasingly unstable, unless R is a bulky substituent group (*12*). This instability probably originates from intermolecular polymerization of linear carbon, like the 1,4-addition polymerization shown in the first figure. Therefore, one possible solution is to use a matrix to isolate chains of linear carbon. Hlavatý *et al.* (*2*) used this approach in their effort to convert C_4I_2 in a microporous silica to linear carbon using ultraviolet irradiation and were successful in producing at least short carbon chain segments. Linear carbon chains of different lengths are stable when trapped at low temperature in a solid argon matrix (*13*). Also, there are reports (*14*, *15*) that linear carbon occasionally forms as highly stable species inside carbon multiwalled nanotubes during carbon arc synthesis. Perhaps almost all nanotubes in a sample could be filled by linear carbon by loading them with C_4I_2, polymerizing the C_4I_2, and then extruding the iodine from the poly(C_4I_2).

Does the above instability mean that it will be forever impossible to obtain even very small crystals of linear carbon that are stable at room temperature? Perhaps not, given that the addition polymerization of diacetylenes is thermodynamically uphill until three or four molecules react together. Hence, nanofibers composing about seven bundled chains of linear carbon might be sufficiently stable for applications.

It might even be possible to make use of the high strength and high modulus of suitably isolated chains of linear carbon for polymer composites. Another possibility is to use single chains of linear carbon as an extremely narrow (3.4 Å wide) channel for field-effect transistors. Unlike the presently available single-walled carbon nanotubes, all linear carbon chains would be identical and semiconducting, thereby eliminating the need for sorting through a pile of nanotubes to make a device.

In fact, the technology for polymerizing individual polydiacetylene chains of desired length on a substrate already exists (*16*) and seems highly applicable for either linear arrays of C_4I_2 molecules or monomer complexes like those described by Sun *et al*. It might be possible to convert these individual poly(C_4I_2) chains to chains of linear carbon by electron-beam irradiation, photolysis, heating, or exposure to an agent facilitating iodine extrusion, such as an evaporated alkali metal. Whether or not linear carbon synthesis is practical by the poly(C_4I_2) route, Sun *et al*. have provided both an exciting new electronic material and a route that could lead to novel related and derivative materials.

Polymer crystal. Recent photograph of a 30-year-old, 4-cm-long polydiacetylene sample with substituent $(CH_2)_3OCONHC_2H_5$. To make this crystal, a diacetylene monomer crystal was grown from its melt, and γ-ray polymerization was used to quantitatively convert the monomer crystal to the polydiacetylene crystal.

CREDIT: S. FANG, NANOTECH INSTITUTE/UNIVERSITY OF TEXAS AT DALLAS

References and Notes

1. A. Sun, J. W. Lauher, N. S. Goroff, *Science* **312**, 1030 (2006).
2. J. Hlavatý, J. Rathousky, A. Zukal, L. Kavan, Carbon **39**, 53 (2001).
3. N. S. Goroff, S. M. Curtis, J. A. Webb, F. W. Fowler, J. W. Lauher, *Org. Lett.* **7**, 1891 (2005).
4. S. M. Curtis, N. Lee, F. W. Fowler, J. W. Lauher, *Cryst. Growth Des.* **5**, 2313 (2005).
5. G. Wegner, *Z. Naturforsch.* B **24**, 824 (1969).
6. R. H. Baughman, R. R. Chance, *Ann. N.Y. Acad. Sci.* **313**, 705 (1978).
7. D. Bloor, R. R.Chance, Eds., Polydiacetylenes: Synthesis, Structure, and Electronic Properties, NATO Advanced Study Institutes Series E, vol. 102 (Nijhoff, Dordrecht, Netherlands, 1985).
8. A. Sarkar, S. Okada, H. Matsuzawa, H. Matsuda, H. Nakanishi, *J. Mater. Chem.* **10**, 819 (2000).
9. G. N. Patel, A. F. Preziosi, R. H. Baughman, U.S. Patent 3,999,946 (1976).
10. An article by PATH (a nonprofit international health organization) on vaccine vial time-temperature indicators based on the color changes associated with diacetylene polymerization can be found at www.path.org/publications/pub.php?id=1135.
11. R. B. Heimann, S. E. Evsyukov, L. Kavan, Eds., *Carbyne and Carbynoid Structures* (Kluwer, Dordrecht, Netherlands, 1999).
12. T. Gibtner, F. Hampel, J.-P. Gisselbrecht, A. Hirsch, *Chem. Eur. J.* **8**, 408 (2002).
13. K. Szczepanski, J. Fuller, S. Ekern, M. Vala, *Spectrochim. Acta A* **57**, 775 (2001).
14. Z, Wang *et al.*, *Phys. Rev. B* **61**, R2472 (2000).
15. X. Zhao *et al. Phys. Rev. Lett.* **90**, 187401 (2003).
16. Y. Okawa, M. Aono, *Nature* **409**, 683 (2001).

10.1126/science.1125999

Genomics and the Irreducible Nature of Eukaryote Cells

C. G. Kurland,[1] L. J. Collins,[2] D. Penny[2]*

Large-scale comparative genomics in harness with proteomics has substantiated fundamental features of eukaryote cellular evolution. The evolutionary trajectory of modern eukaryotes is distinct from that of prokaryotes. Data from many sources give no direct evidence that eukaryotes evolved by genome fusion between archaea and bacteria. Comparative genomics shows that, under certain ecological settings, sequence loss and cellular simplification are common modes of evolution. Subcellular architecture of eukaryote cells is in part a physical-chemical consequence of molecular crowding; subcellular compartmentation with specialized proteomes is required for the efficient functioning of proteins.

Comparative genomics and proteomics have strengthened the view that modern eukaryote and prokaryote cells have long followed separate evolutionary trajectories. Because their cells appear simpler, prokaryotes have traditionally been considered ancestors of eukaryotes (*1*–*4*). Nevertheless, comparative genomics has confirmed a lesson from paleontology: Evolution does not proceed monotonically from the simpler to the more complex (*5*–*9*). Here, we review recent data from proteomics and genome sequences suggesting that eukaryotes are a unique primordial lineage.

Mitochondria, mitosomes, and hydrogenosomes are a related family of organelles that distinguish eukaryotes from all prokaryotes (*10*). Recent analyses also suggest that early eukaryotes had many introns (*11*, *12*), and RNAs and proteins found in modern spliceosomes (*13*). Indeed, it seems that life-history parameters affect intron numbers (*14*, *15*). In addition, "molecular crowding" is now recognized as an important physical-chemical factor contributing to the compartmentation of even the earliest eukaryote cells (*16*, *17*).

Nuclei, nucleoli, Golgi apparatus, centrioles, and endoplasmic reticulum are examples of cellular signature structures (CSSs) that distinguish eukaryote cells from archaea and bacteria. Comparative genomics, aided by proteomics of CSSs such as the mitochondria (*18*, *19*), nucleoli (*20*, *21*), and spliceosomes (*13*, *22*), reveals hundreds of proteins with no orthologs evident in the genomes of prokaryotes; these are the eukaryotic signature proteins (ESPs) (*23*, *24*). The many ESPs within the subcellular structures of eukaryote cells provide landmarks to track the trajectory of eukaryote genomes from their origins. In contrast, hypotheses that attribute eukaryote origins to genome fusion between archaea and bacteria (*25*–*30*) are surprisingly uninformative about the emergence of the cellular and genomic signatures of eukaryotes (CSSs and ESPs). The failure of genome fusion to directly explain any characteristic feature of the eukaryote cell is a critical starting point for studying eukaryote origins.

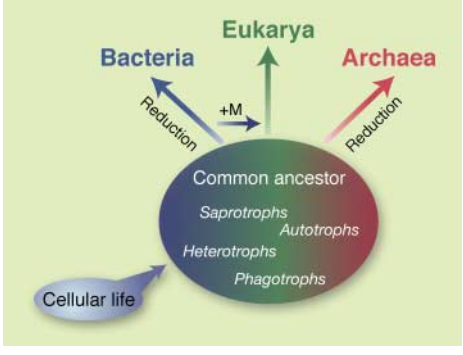


Fig. 1. The common ancestor of eukaryotes, bacteria, and archaea may have been a community of organisms containing the following: autotrophs that produced organic compounds from CO_2 either photosynthetically or by inorganic chemical reactions; heterotrophs that obtained organics by leakage from other organisms; saprotrophs that absorbed nutrients from decaying organisms; and phagotrophs that were sufficiently complex to envelop and digest prey. +M: endosymbiosis of mitochondrial ancestor.

It is agreed that, whether using gene content, protein-fold families, or RNA sequences (*31*–*36*), the unrooted tree of life divides into archaea, bacteria, and eukaryotes (Fig. 1). On such unrooted trees, the three domains diverge from a population that can be called the last universal common ancestor (LUCA). However, LUCA (*37*) means different things to different people, so we prefer to call it a common ancestor; in this case it is the hypothetical node at which the three domains coalesce in unrooted trees.

There are links between comparative genomics and the ecology of organisms. These include the aerobic/anaerobic states of the environment and the adaptive fit of organelles such as mitochondria, hydrogenosomes, and mitosomes (*10*, *18*, *19*, *38*–*41*). In addition to the advantages from oxidative metabolism and/or oxygen detoxification, other advantages must have accrued from having a cellular compartment with dense proteomes (*15*, *38*, *42*). Ecological specialization can account for the differences between prokaryote and eukaryote cell architectures and genome sizes. Small prokaryote cells with streamlined genomes may reflect adaptation to rapid growth and/or minimal resource use by autotrophs, heterotrophs, and saprotrophs. Divergent evolutionary paths may emerge with the adoption of a phagotrophic-feeding mode in an ancestor of eukaryotes. This uniquely eukaryote feeding mode requires a larger and more complex cell, consistent with earlier suggestions that a unicellular raptor (predator), which acquired a bacterial endosymbiont/mitochondria lineage, became the common ancestor of all modern eukaryotes (*3*, *4*, *43*). Indeed, predator/prey relationships may provide the ecological setting for the divergence of the distinctive cell types adopted by eukaryotes, bacteria, and archaea.

Proteomics of Cell Compartments

Comparative genomics and proteomics reveal phylogenetic relationships between proteins making up eukaryote subcellular features and those found in prokaryotes. We distinguish three main phylogenetic classes; the first are proteins that are unique to eukaryotes: the ESPs. The ESPs we place in three subclasses: proteins arising de novo in eukaryotes; proteins so divergent to homologs of other domains that their relationship is largely lost; or finally, descendants of proteins that are lost from other domains, surviving only as ESPs in eukaryotes.

The second class contains interdomain horizontal gene transfers; these are proteins occurring in two domains with the lineage of one domain rooted within their homologs in a second domain (*44*). The third class contains homologs found in at least two domains, but the proteins of one domain are not rooted within another domain(s); instead, the homologs appear to descend from the common ancestor (Fig. 1). Most eukaryote proteins shared by prokaryotes are distant, rather than close, relatives. Thus, proteins shared between domains appear to be descendants of the common ancestor; few seem to result from interdomain lateral gene transfer (*31*–*35*).

Although the genomes of mitochondria are clearly descendants of α-proteobacteria (*45*, *46*), proteomics and comparative genomics identify relatively few proteins in yeast and human

[1]Department of Microbial Ecology, Lund University, Lund, Sweden. [2]Allan Wilson Center for Molecular Ecology and Evolution, Massey University, Palmerston North, New Zealand.

*To whom correspondence should be addressed. E-mail: d.penny@massey.ac.nz

mitochondria descended from the ancestral bacterium (*17*, *18*, *36*, *47*). Several hundred genes have been transferred from the ancestral bacterium to the nuclear genome, but most proteins from the original endosymbiont have been lost. For yeast, the largest protein class contains more than 200 eukaryote proteins (ESPs) targeted to the mitochondrion but encoded in the nucleus. In addition, the yeast nucleus encodes 150 mitochondrial proteins not uniquely identifiable with a single domain but apparently eukaryotic descendants from the common ancestor. Accordingly, the yeast and human mitochondria proteomes emerge largely as products of the eukaryotic nuclear genome (85%) and only to a lesser degree (15%) as direct descendants of endosymbionts (*17*, *18*, *36*, *45*). The strong representation of ESPs in their proteomes means that mitochondria and their descendants are usefully viewed as "honorary" CSSs.

There are substantial numbers of ESPs in the other CSSs. For the proteome of the reduced anaerobic parasite *Giardia lamblia* (*23*), searches of 2136 proteins found in each of *Saccharomyces cerevisiae*, *Drosophila melanogaster*, *Caenorhabditis elegans*, and *Arabidopsis thaliana* yielded 347 ESPs for *G. lamblia*. This was reduced to roughly 300 by rigorous screening, with ESPs distributed between nuclear and cytoplasmic compartments (Fig. 2) (*48*). The ubiquity of the ESPs and the absence of archaeal descendants are not easily explained by a prokaryote genome fusion model (*49*). The simplest interpretation is that the host for the endosymbiont/mitochondrial lineage was an ancestral eukaryote.

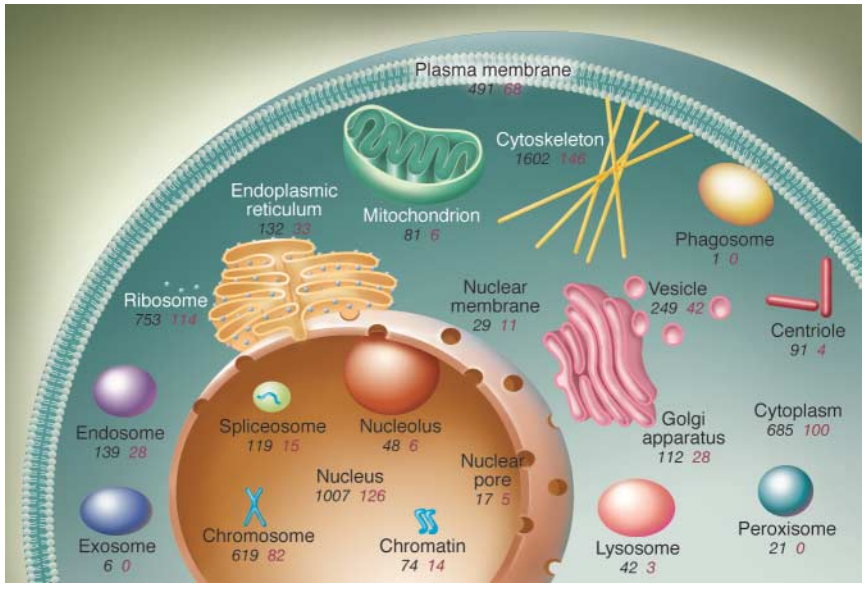


Fig. 2. Distribution of ESPs in the proteome of *G. lamblia*. ESPs (*23*) were matched to the human International Protein Index data set (*48*) and then assigned to individual CSSs based on their gene ontology annotations. A protein may be present in more than one CSS (e.g., a protein involved in transport from the nucleus to the cytoplasm will be assigned to both CSSs). Black numbers are the number of proteins assigned to each CSS from the total *G. lamblia* proteome (AACB00000000) (3077 ORFs matched and linked to gene ontology); red numbers are the ESPs assigned to each CSS (320 proteins matched and linked to gene ontology).

Similar results are obtained for another reduced eukaryote, the intracellular parasite *Encephalitozoon cuniculi*. A recent study (*24*) identified 401 ESPs, of which 295 had homologs among the ESPs of *G. lamblia* (*23*). Two major categories of ESPs in the *G. lamblia* and *E. cuniculi* genomes were distinguished: those associated with the CSSs (Fig. 2) and those involved in control functions such as guanosine triphosphate (GTP) binding proteins, kinases, and phosphatases (*7*). It was also observed (*23*) that many characteristic eukaryotic proteins with weak sequence homology to prokaryotic proteins but more convincing homologies of structural fold such as the actins, tubulins, kinesins, ubiquitins, and some GTP binding proteins are among the most highly conserved eukaryotic proteins. These may be descendants of the common ancestor recruited early in the evolution of the eukaryotic nuclear genome.

Nucleolar proteomes (*20*, *21*) are examples of essential eukaryote compartments not wrapped in double membranes and where there is no suspicion of an endosymbiotic origin. From 271 proteins in the human nucleolar proteome, 206 protein folds were identified and classified phylogenetically (*20*, *21*). Of these, 109 are eukaryotic signature folds, and the remaining ones appear to be descendants of the common ancestor, occurring in two or three domains.

The spliceosome is a unique molecular machine that removes introns from eukaryote mRNAs (*22*). Even though we do not know the ancestral processing signals for the earliest eukaryotes (*50*), roughly half of the 78 spliceosomal proteins likely to be present in the ancestral spliceosome are ESPs, (*13*) whereas the other half containing the Sm/LSm proteins (*51*) have homologs in bacteria and archaea (*13*). These distributions of both ESPs as well as of putative descendants of the common ancestor suggest that many components of modern spliceosomes were present in the common ancestor (*52*).

The subdivision into subcellular compartments (CSSs) with characteristic proteomes restricts proteins to volumes considerably smaller than the whole cell. Concentrations of macromolecules in cells are very high, typically between 20 and 30% of weight or volume (*53*). Such densities are described as "molecular crowding" because the space between macromolecules is much less than their diameters; consequently, diffusion of proteins in cells is retarded (*54*). Molecular crowding favors macromolecular associations, large complexes, and networks of proteins that support biological functions (*16*, *17*, *53*).

High densities enhance the association kinetics of small molecules with proteins because the excluded volumes of the proteins reduce the effective volume through which small molecules diffuse (*55*). The sum of these effects is that the high macromolecular densities within CSSs enhance the kinetic efficiencies of proteins. The same principles apply to the smaller prokaryotic cells, but the effects are accentuated in larger cells. Subdividing high densities of proteins into more or less distinct compartments containing functionally interactive macromolecules is expected to be an early feature of the eukaryote lineage. The distinctive proteome of nucleoli demonstrates that compartmentation does not require an enclosing membrane. Furthermore, cell fusion is not required to account for, nor does it explain (*49*), the large number of eukaryote cell compartments.

Selection Gives and Selection Takes

Genomes evolve continuously through the interplay of unceasing mutation, unremitting competition, and ever-changing environments. Both sequence loss and sequence gain can result. In general, expanded genome size, along with augmented gene expression, increases the costs of cell propagation so the evolution of larger genomes and larger cells requires gains in fitness that compensate (*15*, *56*, *57*). Conversely, genome reduction is expected to lower the costs of propagation. There is an ever-present potential to improve the efficiency of cell propagation by reductive evolution.

Environmental shifts may neutralize sequences, leaving no selective pressure to maintain them against the persistent flux of deleterious mutations. Such neutralized sequences eventually and inevitably disappear because of "mutational meltdown" (*14*, *15*, *56*, *57*). Genome reduction can be achieved through differential loss of coding and noncoding sequences (com-

paction) (*57*). *Theileria* has evolved through gene loss as well as compaction of its intergenic spaces, whereas *Paramecium* has eliminated only a small length of genes but markedly reduced the number of its introns (*57*). The complex genomes of some vertebrates (pufferfish, *Takifugu*) are so highly compacted that their genome lengths are reduced to one-eighth that of other vertebrates (*58*). Extreme cellular simplification is observed among anaerobic protists, including simplification of CSSs such as mitochondria and the Golgi apparatus (*59*–*64*). *S. cerevisiae*, which underwent a whole-genome duplication, subsequently purged ~85% of the duplicated sequences (*65*, *66*). The evolution of genome content is clearly not monotonic (Fig. 3) (*67*, *68*). Genome sizes on the branches of a phylogenetic tree of fungi show irregular genome enlargement (including duplication) and reduction. Examples of ecological circumstances driving genome reduction are seen in many intracellular endosymbionts and parasites, which gain few genes but lose many genes responsible for metabolic flexibility (*6*–*8*, *69*).

The mitochondrion is even more extreme in its reductive evolution; its ancestral bacterial genome has been reduced to a vestigial microgenome supported by a predominantly eukaryote proteome (*18*, *19*). Genomes of modern mitochondria encode between 3 and 67 proteins (*44*), whereas the smallest known free-living α-proteobacterium (*Bartonella quintana*) encodes ~1100 proteins (*70*). Taking *Bartonella* as a minimal genome for the free-living ancestor of mitochondria, nearly all of the bacterial coding sequences have been lost from the organelle, though not necessarily from the eukaryote cell. The mitochondrial genome of the protist *Reclinomonas americana* is the largest known but has still lost more than 95% of its original coding capacity.

This abbreviated account of genome reduction illustrates the Darwinian view of evolution as a reversible process in the sense that "eyes can be acquired and eyes can be lost." Genome evolution is a two-way street. This bidirectional sense of reversibility is important as an alternative to the view of evolution as a rigidly monotonic progression from simple to more complex states, a view with roots in the 18th-century theory of orthogenesis (*71*). Unfortunately, such a model has been tacitly favored by molecular biologists who appeared to view evolution as an irreversible march from simple prokaryotes to complex eukaryotes, from unicellular to multicellular. The many well-documented instances of genome reduction provide a necessary corrective measure to the often-unstated assumption that eukaryotes must have originated from prokaryotes.

The Hunt for the Phagotrophic Unicellular Raptor

Proteomics, together with comparative genomics, allows glimpses of the cell structure of eukaryote ancestors. They are likely to have had introns as well as the complex machinery for removing them, and much of that RNA processing machinery still exists in their descendants (*13*, *22*, *51*). Because of molecular crowding, it is expected that interacting proteins would tend to accumulate in functional domains, making rudimentary CSSs early features of the large-celled eukaryotes. We cannot say whether there was a substantial period of time after the emergence of cells when there were no unicellular raptors or predators—a Garden of Eden. However, the identification among prokaryotes of orthologs with structural affinities to actins, tubulins, kinesins, and ubiquitins (*72*, *73*) is consistent with some early organisms having evolved a phagotrophic life-style. This echoes a recurrent theme (*3*, *4*, *43*) in which it was supposed that the earliest eukaryotes could feed as unicellular "raptors."

We expect that the earliest organisms were primarily auxotrophs, heterotrophs, and saprotrophs—an excellent community to support raptors. Phagotrophy is a hallmark of eukaryotic cells and is unknown among modern prokaryotes, and so it is natural to reconsider this feeding mode as a defining feature of ancestral eukaryotes. Cavalier-Smith (*43*) suggested that the ancestors of eukaryotes were phagotrophic, anaerobic free-living protists, called archeozoa. He also identified present-day anaerobic parasites such as *Entamoeba*, *Giardia*, and Microsporidia as archeozoa. However, these organisms are descendants of aerobic, mitochondriate eukaryotes (*10*). Genome reduction and cellular simplification are hallmarks of parasites and symbionts (*6*–*8*, *46*, *69*). Indeed, most of the eukaryotic anaerobes studied so far are parasites or symbionts of multicellular creatures.

For the reasons outlined above, we favor the idea (*3*, *4*) that the host that acquired the mitochondrial endosymbiont was a unicellular eukaryote predator, a raptor. The emergence of unicellular raptors would have had a major ecological impact on the evolution of the gentler descendants of the common ancestor. These may have responded with several adaptive strategies: They might outproduce the raptors by rapid growth or hide from raptors by adapting to extreme environments. Thus, the hypothetical eukaryote raptors may have driven the evolution of their autotrophic, heterotrophic, and saprotrophic cousins in a reductive mode that put a premium on the relatively fast-growing, streamlined cell types we call prokaryotes (*74*).

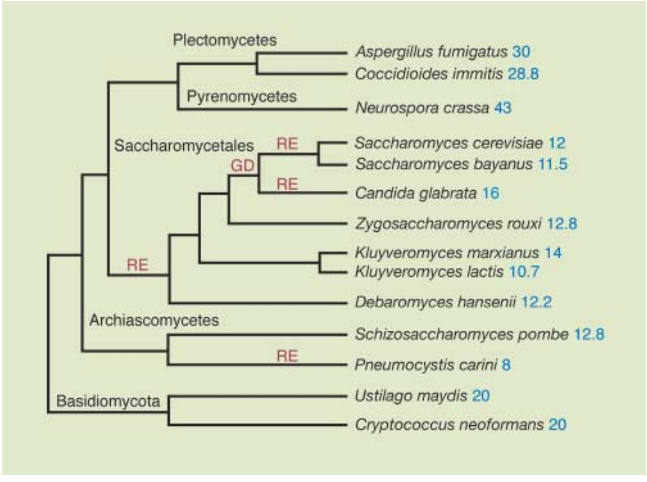


Fig. 3. Genome sizes (in megabases) can increase and decrease in lineages because of events such as genome duplication and reductive evolution, as illustrated in this fungal phylogeny [adapted from (*67*, *68*)]. Genome sizes were obtained from the National Center for Biotechnology Information (NCBI) Genome biology (www.ncbi.nih.gov/Genomes/) database. GD, genome duplication; RE, reductive evolution.

Concluding Remarks

Genomics and proteomics have greatly increased our awareness of the uniqueness of eukaryote cells. This, together with increased understanding of molecular crowding, as well as the dynamic, often reductive nature of genome evolution, offers a new view of the origin of eukaryote cells. The eukaryotic CSSs define a unique cell type that cannot be deconstructed into features inherited directly from archaea and bacteria. Only a small fraction (~15%) of α-proteobacterial proteins are identified in the yeast and human mitochondrial proteomes; none seem to be direct descendants of archaea, and roughly half seem to be exclusively eukaryotic (*18*, *19*, *38*, *47*). The identification of the α-proteobacterial descendants in this proteome validates the phylogenetic distinction between direct descent from genes transferred to the host from the bacterial endosymbiont, as opposed to descent from a hypothetical common ancestor.

ESPs are important markers of the novel evolutionary trajectory of modern eukaryotes. In contrast, most proteins occur in more than one domain (*31*–*36*), and most of these could derive from the common ancestor. We take the relative abundance of signature proteins among eukaryotes to indicate that their genomes typically have a greater coding capacity than those of prokaryotes. It remains to be seen which ESPs have been lost from prokaryotes and which have been acquired by eukaryotes during their evolution.

The hypothetical fusion of an archaeon and a bacterium explains nothing about the special features of the modern eukaryote cell (*49*), nor the many signature proteins. Nothing in global phylogenies based on ribosomal RNA, pooled proteins, and protein-fold families indicates that genome fusion generated the eukaryote lineage. Perhaps interest in fusion models arose because BLAST searches suggest that different eukaryotic coding sequences are sometimes more closely related to archaeal homologs and other times more closely related to bacterial homologs (*49*). These weak domain-specific affinities do need to be understood and alternative explanations found. However, in our view (*49*), they do not indicate that the eukaryote genome arose as a mosaic pieced together from archaeal and bacterial genomes.

It is an attractively simple idea that a primitive eukaryote took up the endosymbiont/mitochondrion by phagocytosis (*3*, *4*, *43*). A unicellular raptor with a larger, more complex cell structure than that of present-day prokaryotes is envisioned as the host of the ancestral endosymbiont. This scenario, which is not contradicted by new data derived from comparative genomics and proteomics, is a suitable starting point for future work. Acquisition of genome sequences from free-living eukaryotes among basal lineages is a high priority.

References and Notes

1. A. H. Knoll, *Science* **256**, 622 (1992).
2. S. L. Baldauf, *Science* **300**, 1703 (2003).
3. R. Y. Stanier, C. B. van Niel, *Arch. Mikrobiol.* **42**, 17 (1962).
4. C. de Duve, *Ann. N. Y. Acad. Sci.* **386**, 1 (1982).
5. G. G. Simpson, *Tempo and Mode in Evolution* (Columbia Univ. Press, New York, 1944).
6. S. G. Andersson, C. G. Kurland, *Trends Microbiol.* **6**, 263 (1998).
7. L. Klasson, S. G. E. Andersson, *Trends Microbiol.* **12**, 37 (2004).
8. M. V. Olsen, *Am. J. Hum. Genet.* **64**, 18 (1999).
9. P. Forterre, H. Philippe, *Bioessays* **21**, 871 (1999).
10. T. M. Embley, W. Martin, *Nature* **440**, 623 (2006).
11. S. W. Roy, W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 5773 (2005).
12. F. Raible *et al.*, *Science* **310**, 1325 (2005).
13. L. J. Collins, D. Penny, *Mol. Biol. Evol.* **22**, 1053 (2005).
14. M. Lynch, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 6118 (2002).
15. D. C. Jeffares, T. Mourier, D. Penny, *Trends Genet.* **22**, 16 (2006).
16. G. J. Pielak, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 5901 (2005).
17. R. J. Ellis, *Curr. Opin. Struct. Biol.* **11**, 114 (2001).
18. O. Karlberg, B. Canback, C. G. Kurland, S. G. Andersson, *Yeast* **17**, 170 (2000).
19. C. G. Kurland, S. G. Andersson, *Microbiol. Mol. Biol. Rev.* **64**, 786 (2000).
20. J. S. Andersen *et al.*, *Nature* **433**, 77 (2005).
21. E. Staub, P. Fiziev, A. Rosenthal, B. Hinzmann, *Bioessays* **26**, 567 (2004).
22. M. S. Jurica, M. J. Moore, *Mol. Cell* **12**, 5 (2003).
23. H. Hartman, A. Fedorov, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 1420 (2002).
24. A. Fedorov, H. Hartman, *J. Mol. Evol.* **59**, 695 (2004).
25. J. P. Gogarten, L. Olendzenski, E. Hilario, C. Simon, K. E. Holsinger, *Science* **274**, 1750 (1996).
26. D. F. Feng, G. Cho, R. F. Doolittle, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 13028 (1997).
27. S. Ribeiro, G. B. Golding, *Mol. Biol. Evol.* **15**, 779 (1998).
28. P. Lopez-Garcia, D. Moreira, *Trends Biochem. Sci.* **24**, 88 (1999).
29. T. Cavalier-Smith, *Int. J. Syst. Evol. Microbiol.* **52**, 297 (2002).
30. M. C. Rivera, J. A. Lake, *Nature* **431**, 152 (2004).
31. C. R. Woese, O. Kandler, M. L. Wheelis, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4576 (1990).
32. J. O. Korbel, B. Snel, M. A. Huynen, P. Bork, *Trends Genet.* **18**, 158 (2002).
33. B. Snel, P. Bork, M. A. Huynen, *Genome Res.* **12**, 17 (2002).
34. R. F. Doolittle, *Annu. Rev. Biochem.* **64**, 287 (1995).
35. G. Caetano-Anolles, D. Caetano-Anolles, *Genome Res.* **13**, 1563 (2003).
36. S. Yang, R. F. Doolittle, P. E. Bourne, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 373 (2005).
37. D. Penny, A. Poole, *Curr. Opin. Genet. Dev.* **9**, 672 (1999).
38. T. Gabaldon, M. A. Huynen, *Science* **301**, 609 (2003).
39. O. G. Berg, C. G. Kurland, *Mol. Biol. Evol.* **17**, 951 (2000).
40. T. M. Embley *et al.*, *IUBMB Life* **55**, 387 (2003).
41. B. A. Williams, R. P. Hirt, J. M. Lucocq, T. M. Embley, *Nature* **418**, 865 (2002).
42. A. Regoes *et al.*, *J. Biol. Chem.* **280**, 30557 (2005).
43. T. Cavalier-Smith, *Nature* **326**, 332 (1987).
44. B. Canback, S. G. Andersson, C. G. Kurland, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 6097 (2002).
45. S. G. Andersson *et al.*, *Nature* **396**, 133 (1998).
46. M. W. Gray, G. Burger, B. F. Lang, *Science* **283**, 1476 (1999).
47. T. Gabaldon, M. A. Huynen, *Bioinformatics* **21** (suppl. 2), ii144 (2005).
48. P. J. Kersey *et al.*, *Proteomics* **4**, 1985 (2004).
49. See Supporting Online Material.
50. L. J. Collins, D. Penny, *Mol. Biol. Evol.* **23**, 901 (2006).
51. J. D. Beggs, *Biochem. Soc. Trans.* **33**, 433 (2005).
52. A. Poole, D. Jeffares, D. Penny, *Bioessays* **21**, 880 (1999).
53. R. J. Ellis, *Trends Biochem. Sci.* **26**, 597 (2001).
54. K. Luby-Phelps, *Int. Rev. Cytol.* **192**, 189 (2000).
55. M. Ehrenberg, C. G. Kurland, *Q. Rev. Biophys.* **17**, 45 (1984).
56. O. G. Berg, C. G. Kurland, *Mol. Biol. Evol.* **19**, 2265 (2002).
57. P. J. Keeling, C. H. Slamovits, *Curr. Opin. Genet. Dev.* **15**, 601 (2005).
58. S. Aparicio *et al.*, *Science* **297**, 1301 (2002).
59. J. B. Dacks *et al.*, *Proc. Biol. Sci.* **270** (suppl. 2), S168 (2003).
60. C. G. Clark, A. J. Roger, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6518 (1995).
61. A. Germot, H. Philippe, H. Le Guyader, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14614 (1996).
62. J. Tovar *et al.*, *Nature* **426**, 172 (2003).
63. A. Akhmanova *et al.*, *Mol. Microbiol.* **30**, 1017 (1998).
64. B. Boxma *et al.*, *Nature* **434**, 74 (2005).
65. K. H. Wolfe, D. C. Shields, *Nature* **387**, 708 (1997).
66. B. Dujon *et al.*, *Nature* **430**, 35 (2004).
67. J. Cai, I. N. Roberts, M. D. Collins, *Int. J. Syst. Bacteriol.* **46**, 542 (1996).
68. A. C. Padovan, G. F. Sanson, A. Brunstein, M. R. Briones, *J. Mol. Evol.* **60**, 726 (2005).
69. N. A. Moran, *Cell* **108**, 583 (2002).
70. C. M. Alsmark *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 9716 (2004).
71. P. J. Bowler, *The Non-Darwinian Revolution: Reinterpreting a Historical Myth* (Johns Hopkins Univ. Press, Baltimore, 1988).
72. R. R. Copley, J. Schultz, C. P. Ponting, P. Bork, *Curr. Opin. Struct. Biol.* **9**, 408 (1999).
73. C. P. Ponting, L. Aravind, J. Schultz, P. Bork, E. V. Koonin, *J. Mol. Biol.* **289**, 729 (1999).
74. J. E. Darnell Jr., *Science* **202**, 1257 (1978).

75. We thank G. Jameson and M. Oliveberg for very useful discussions and help with the literature. Supported by The New Zealand Marsden Fund (L.J.C.) and The Royal Physiographic Society, Lund (C.G.K.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/312/5776/1011/DC1
SOM Text
Figs. S1 to S4
References

10.1126/science.1121674

Dispersal Limitations Matter for Microbial Morphospecies

Richard J. Telford,[1,2]* Vigdis Vandvik,[2] H. J. B. Birks[1,2,3]

Geographic range restrictions are typical for macroorganisms. In contrast, many microorganisms appear to have cosmopolitan distributions, an observation Baas Becking formulated as "everything is everywhere, but, the environment selects" (*1*). Finlay (*2*) argues that microbial biogeography is fundamentally different, because vast population sizes drive ubiquitous dispersal. This hypothesis has been challenged by evidence of dispersal limitations. Taxonomists argue that they find endemic morphospecies—for example, in the Yellowstone Lake diatom flora (*3*)—and genetic data demonstrate that microbial organisms can exhibit classical biogeographical patterns, such as provincialism and distance decay (*4*). Unfortunately, taxonomic uncertainties, undersampling, and different interpretations of the genetic variability frustrate unequivocal testing of the ubiquitous dispersal hypothesis with these types of evidence (*2*, *4*, *5*).

Understanding the rate of global dispersal is key to discriminating between these alternative views of microbial biogeography. Is dispersal so rapid that regional biota are homogenized into a single global metacommunity, or is it slow enough to allow signatures of regional-scale processes to develop? We address this question by exploring regional species richness-environment relationships. If dispersal is ubiquitous, all regions should share one underlying richness-environment relationship, governed by the global species pool. If dispersal is limited, regional-scale metacommunity processes (*6*), such as colonization-extinction dynamics, will decouple regional species pools from the global pool, allowing richness-environment relationships to vary between regions. We predict that, if dispersal is limited, regional richness relationships will depend on regional habitat availability (Fig. 1A).

We tested this prediction by investigating lake-diatom species richness along pH gradients in 16 regional data sets from Europe and North America. Diatoms are a pH-sensitive, species-rich group of microscopic siliceous algae, common in freshwater lakes. By concentrating on richness patterns (Fig. 1), our analysis does not require taxonomic consistency between data sets, only that there are no pH-related biases in the richness estimates.

The species richness-pH relationships differ markedly among regions (Fig. 1B and fig. S1), and richness optima, which vary by 2.5 pH units, are strongly correlated with the regionally most common environments (Fig. 1C). Adjacent regions (e.g., within Fennoscandia) can have very different richness pH-optima, whereas latitudinal trends are weak (fig. S2). This suggests that diatom metacommunities are regional rather than global in scale. Despite the enormous populations and high passive dispersal potential of diatoms, their dispersal at the global scale must be too slow to override regional metacommunity processes.

These results complement recent genetic work by demonstrating that dispersal limitation can affect microbial community assembly, even at the relatively coarse taxonomic resolution of morphospecies. Our results help reconcile the positions of the adherents and detractors of the ubiquitous dispersal hypothesis. Although many microbial species may indeed have a cosmopolitan distribution, it is probably attained slowly and incrementally. This allows time for regional genetic differences to develop and for endemic taxa specializing in rare or isolated habitats (*3*) to evolve as their probability of dispersal to suitable distant sites approaches zero. Dispersal limitations, which are so important in shaping the diversity and distributions of macroorganisms (*6*), also play a critical role in the microbial world.

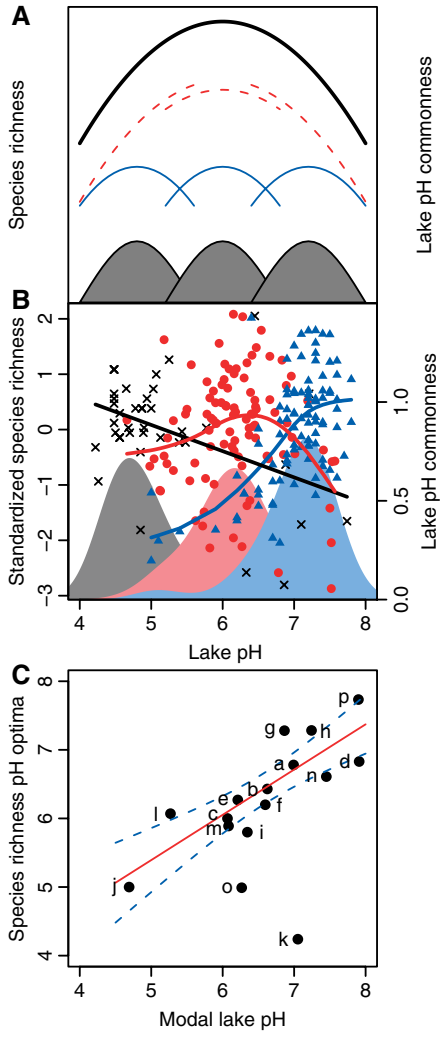


Fig. 1. Predicted patterns of lake diatom species richness in relation to pH (**A**) for three regions with different pH availability (gray areas) if metacommunities operate at the regional (solid blue) or global (dashed red) scale. The thick black line shows a hypothetical global species richness. Observed lake pH availability and standardized rarefied species richness in individual lakes (**B**) for North Florida (black), South Norway (red), and Finnish Lapland (blue). Thick lines show trends in species richness estimated with generalized additive models. Species richness pH optima against modal lake pH (**C**) in 16 regional data sets (*7*) from Europe (a to i) and North America (j to p). Regions are labeled from south to north in each continent. The modal lake pH is estimated from the pH density. Species richness pH optima are estimated using LOWESS (span = ⅔). If richness reached a plateau rather than an optimum, the break of slope was used instead. The linear regression of richness optima on most common lake pH (solid line with dashed confidence interval) is highly significant ($P < 0.001$; excluding two outliers, regions o and k).

References and Notes

1. R. de Wit, T. Bouvier, *Environ. Microbiol.* **8**, 755 (2006).
2. B. J. Finlay, *Science* **296**, 1061 (2002).
3. E. C. Theriot, S. C. Fritz, C. Whitlock, D. J. Conley, *Paleobiology* **32**, 38 (2006).
4. R. J. Whitaker, D. W. Grogan, J. W. Taylor, *Science* **301**, 976 (2003).
5. T. Fenchel, *Science* **301**, 925 (2003).
6. M. A. Leibold *et al.*, *Ecol. Lett.* **7**, 601 (2004).
7. Materials and methods are available as supporting material on *Science* Online.

8. We thank the diatomists who provided the data. Supported by the Research Council of Norway project NORPLAST-2, Nordic Council of Ministers project DEFINE, and European Union Framework 5 project PACLIVA. This is publication no. A129 from the Bjerknes Centre for Climate Research.

[1]Bjerknes Centre for Climate Research, Allégaten 55, N-5007 Bergen, Norway. [2]Ecological and Environmental Change Research Group, Department of Biology, University of Bergen, Allégaten 41, N-5007 Bergen, Norway. [3]Environmental Change Research Centre, University College London, London, WC1H 0AP, UK.

*To whom correspondence should be addressed. E-mail: richard.telford@bjerknes.uib.no

Supporting Online Material

www.sciencemag.org/cgi/content/full/312/5776/1015/DC1
Materials and Methods
Figs. S1 and S2
Table S1
References

1 February 2006; accepted 27 March 2006
10.1126/science.1125669

Drilling to Gabbro in Intact Ocean Crust

Douglas S. Wilson,[1*†‡§] Damon A. H. Teagle,[2*†‡] Jeffrey C. Alt,[3*‡] Neil R. Banerjee,[4*†‡] Susumu Umino,[5*†] Sumio Miyashita,[6‡] Gary D. Acton,[7*] Ryo Anma,[8‡] Samantha R. Barr,[9*] Akram Belghoul,[10†] Julie Carlut,[11‡] David M. Christie,[12‡] Rosalind M. Coggon,[3*‡] Kari M. Cooper,[7*] Carole Cordier,[13†] Laura Crispini,[14*†] Sedelia Rodriguez Durand,[15†] Florence Einaudi,[10*†] Laura Galli,[16†‡] Yongjun Gao,[17†] Jörg Geldmacher,[18†] Lisa A. Gilbert,[19†] Nicholas W. Hayman,[20‡] Emilio Herrero-Bervera,[21†] Nobuo Hirano,[23‡] Sara Holter,[24†] Stephanie Ingle,[22‡] Shijun Jiang,[25*] Ulrich Kalberkamp,[26*] Marcie Kerneklian,[27*] Jürgen Koepke,[28‡] Christine Laverne,[29*†‡] Haroldo L. Lledo Vasquez,[30†] John Maclennan,[31‡] Sally Morgan,[32‡] Natsuki Neo,[6‡] Holly J. Nichols,[18†] Sung-Hyun Park,[33‡] Marc K. Reichow,[9‡] Tetsuya Sakuyama,[34†] Takashi Sano,[35†] Rachel Sandwell,[36*] Birgit Scheibner,[37‡] Chris E. Smith-Duque,[2†] Stephen A. Swift,[38‡] Paola Tartarotti,[16*†] Anahita A. Tikku,[39‡] Masako Tominaga,[40†‡] Eugenio A. Veloso,[8†‡] Toru Yamasaki,[41‡] Shusaku Yamazaki,[6‡] Christa Ziegler[42*]

Sampling an intact sequence of oceanic crust through lavas, dikes, and gabbros is necessary to advance the understanding of the formation and evolution of crust formed at mid-ocean ridges, but it has been an elusive goal of scientific ocean drilling for decades. Recent drilling in the eastern Pacific Ocean in Hole 1256D reached gabbro within seismic layer 2, 1157 meters into crust formed at a superfast spreading rate. The gabbros are the crystallized melt lenses that formed beneath a mid-ocean ridge. The depth at which gabbro was reached confirms predictions extrapolated from seismic experiments at modern mid-ocean ridges: Melt lenses occur at shallower depths at faster spreading rates. The gabbros intrude metamorphosed sheeted dikes and have compositions similar to the overlying lavas, precluding formation of the cumulate lower oceanic crust from melt lenses so far penetrated by Hole 1256D.

Ocean crust formed at mid-ocean ridges covers more than 60% of Earth's surface, yet our understanding of its accretion at mid-ocean ridges and evolution on the ridge flanks has been severely limited by the extreme difficulty of direct sampling. Remote geophysical measurements have produced longstanding models for the structure of ocean crust, including the size and shape of magma chambers at mid-ocean ridges (*1*–*3*), but the lack of direct sampling of in situ crust has prevented testing these models. Gabbros are coarse-grained mafic rocks commonly formed from slow cooling of magma chambers beneath mid-ocean ridges. Drilling a complete section of upper oceanic crust down to gabbro will enable testing models for the formation and structure of oceanic crust (*4*–*9*).

Multichannel reflection seismic (MCS) profiling of active intermediate and fast-spreading ridges commonly shows bright reflectors at depths of 1 to 4 km that have the properties expected for a thin (20- to 100-m) lens of partial melt (*10*–*14*). These melt lenses extend less than 1 km from the ridge axis and crystallize to form gabbroic rocks. The depth to the reflectors decreases as spreading rate increases (Fig. 1) (*15*, *16*); it is controlled by the rate of magma supply from below and hydrothermal cooling by seawater from above (*7*). Melt lenses are hypothesized to play a critical role in the formation of the lower oceanic crust. According to the "gabbro glacier" model (*6*–*8*), as oceanic crust spreads away from the ridge axis, the accumulated crystal residues in these melt lenses subside to form the lower ocean crust, which is the major portion of the crust. Alternative models, however, argue that the lower crust is formed by injection of sills at various depths (*5*, *17*) and that the geophysically imaged melt lens is simply the most shallow intrusion.

In addition to geophysical studies, our understanding of oceanic basement and particularly the plutonic portion of the crust comes from observations of ancient oceanic rocks exposed on land in ophiolites, seafloor observations of active ridges and deep-sea tectonic exposures, and drilling. The origin of ophiolites in marginal basins and the disruption of tectonically exposed lower crust, however, make the relevance of these observations to intact ocean crust questionable. Previous deep drilling in intact crust has only once penetrated the transition from lavas to dikes, in Ocean Drilling Program (ODP) Hole 504B, which reached a total depth of 1836 m sub-basement (msb) (*18*). Unfortunately, Hole 504B failed to penetrate the dike-gabbro boundary because of hostile drilling conditions in fractured dikes at high temperatures. Although fault-exposed lower ocean crust has been drilled in several places (*19*–*23*), the geological context of such cores is often unknown. The critical transition from dikes to gabbros has previously never been cored.

Deep drilling into basement at Site 1256. Recently, Integrated Ocean Drilling Program (IODP) Expeditions 309 and 312 deepened Hole 1256D in the eastern Pacific to 1507 m below seafloor (mbsf) (1257 msb; msb = mbsf – sediment thickness), drilling through lavas, the underlying sheeted dike complex, and into gabbroic rocks. This is the first penetration of the dike-gabbro boundary in intact ocean crust since the inception of deep sea drilling nearly 40 years ago. Hole 1256D thus provides unique samples of the lithologic transitions in the upper crust, from lavas to dikes and from dikes to gabbros. The dike-gabbro boundary is key to understanding crustal structure and the interplay between magmatic accretion and hydrothermal cooling.

The recognition of an interval of superfast spreading rates, up to 220 mm/year full rate (*24*), on the Cocos-Pacific plate boundary between 19 and 12 million years ago (Ma) led to the choice of ODP Site 1256 (Fig. 2) as the optimal site for deep drilling (*25*). Scientific ocean drilling mainly targets relatively soft, easily cored sediments. In contrast, coring into the underlying, much harder, basaltic basement is less common and most holes are shallow [<300 m (*25*)]. Deep (>500 m) basement drilling requires a substantial commitment of resources but yields major scientific rewards by sampling otherwise inaccessible regions of Earth's interior. A deep drill hole at the fastest possible spreading rate tests the prediction that a melt lens reflector is more shallow at higher spreading rates, and also minimizes the drilling needed to sample an intact section from lavas to gabbros, because the upper crust is thinner. This is an important advantage considering the cost, time, and technical challenges of deep drilling. Assuming that ~300 m of lavas flowed off axis, the depth to gabbros was predicted to be between 1025 to 1300 msb (1275 to 1550 mbsf) at Site 1256 (*26*).

Drilling at Site 1256 was initiated in 2002 on ODP Leg 206 when Hole 1256D was drilled through 250 m of sediment and 502 m into basement (*25*). Coring continued to 1255 mbsf in 2005 by IODP Expedition 309, and recently Expedition 312 deepened the hole to 1507.1 mbsf and into gabbros (*27*). Almost 5 months at Site 1256 were required to achieve the operational and scientific objectives.

Results from drilling. Gabbros were first intersected at 1157 msb (1407 mbsf), within the predicted target zone (*26*). The uppermost crust at Site 1256 is composed of a ~100-m-thick sequence of lava dominated by a single flow up to 75 m thick, requiring at least this much seafloor relief to pool the lava. On modern fast-spreading ridges, such topography does not normally develop until 5 to 10 km from the axis (*28*). The lavas immediately below include

sheet and massive flows, and minor pillow flows. Subvertical, elongate, flow-top fractures filled with quenched glass and hyaloclastite in these lavas indicate flow lobe inflation requiring eruption onto a subhorizontal surface off axis (*29*). Thus, we estimate a total thickness of off-axis lavas of 284 m, close to the assumed thickness. Sheet flows and massive lavas that erupted at the ridge axis make up the remaining extrusive section down to 1004 mbsf, before a lithologic transition is marked by subvertical intrusive contacts and mineralized breccias. Below 1061 mbsf, subvertical intrusive contacts are numerous, indicating the start of a relatively thin, ~350-m-thick, sheeted dike complex that is dominated by massive basalts. Some basalts have doleritic textures, and many are cross-cut by subvertical dikes with common strongly brecciated and mineralized chilled margins. There is no evidence from core or from geophysical wireline logs for substantial tilting of the dikes. This is consistent with seismic reflection images of subhorizontal reflectors in the lower extrusive rocks that are continuous for several kilometers across the site (*30*).

There is a stepwise increase in alteration grade downward from lavas into dikes, with low-temperature phases (<150°C; phyllosilicates and iron oxyhydroxides) in the lavas giving way to dikes partially altered to chlorite and other greenschist minerals (at temperatures >~250°C; Fig. 3). Within the dikes, the alteration intensity and grade increase downward, with actinolite more abundant than chlorite below 1300 mbsf and hornblende present below 1350 mbsf indicating temperatures approaching ~400°C. The dikes have substantially lower porosity (mostly 0.5 to 2%) and higher *P*-wave velocities and thermal conductivity than the lavas; porosity decreases and *P*-wave velocity increases as depth increases in the dikes.

In the lower ~60 m of the sheeted dikes (1348 to 1407 mbsf), basalts are partially to completely recrystallized to distinctive granoblastic textures resulting from contact metamorphism by underlying gabbroic intrusions (Fig. 4). Gabbro and trondhjemite dikes intrude into sheeted dikes at 1407 mbsf, marking the top of the plutonic complex. Two major bodies of gabbro were penetrated beneath this contact, with the 52-m-thick upper gabbro separated from the 24-m-thick lower gabbro by a 24-m screen of granoblastic dikes (Fig. 4). The upper gabbro comprises gabbros, oxide gabbros, quartz-rich oxide diorites, and small trondhjemite dikelets. These rocks are moderately to highly altered by hydrothermal fluids to actinolitic hornblende, secondary plagioclase, epidote, chlorite, prehnite, and laumontite. The relative ferocity of hydrothermal alteration increases with grain size and proximity to intrusive boundaries. The lower gabbro comprises gabbro, oxide gabbro, and subordinate orthopyroxene-bearing gabbro and trondhjemite that are similarly altered, and has clear intrusive contacts with the overlying granoblastic dike screen. Partially resorbed, stoped dike clasts are entrained within both the upper and lower margins of the lower gabbro (Fig. 4G). The lowermost rock recovered from Hole 1256D is a highly altered actinolite-bearing basaltic dike that lacks granoblastic textures, and hence is interpreted to be a late dike that postdates the intrusion of the lower gabbro. Contrary to expectation, porosity increases and *P*-wave velocities decrease stepwise down-

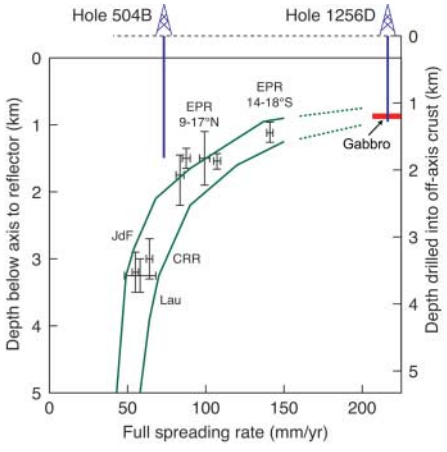


Fig. 1. Depth to axial melt-lens reflector plotted against spreading rate. Depth versus spreading rate predictions from two models of Phipps Morgan and Chen (*7*) are shown, extrapolated subjectively to 200 mm/year (dashed lines). Penetration to date in Holes 504B and 1256D is shown by solid vertical lines, with the depth at which gabbros were intersected indicated in red. According to core descriptions, a thickness of ~300 m of off-axis lavas is shown for Hole 1256D and assumed for Hole 504B. JdF, Juan de Fuca Ridge; Lau, Valu Fa Ridge in Lau Basin; CRR, Costa Rica Rift. [Image modified from (*15*, *16*)]

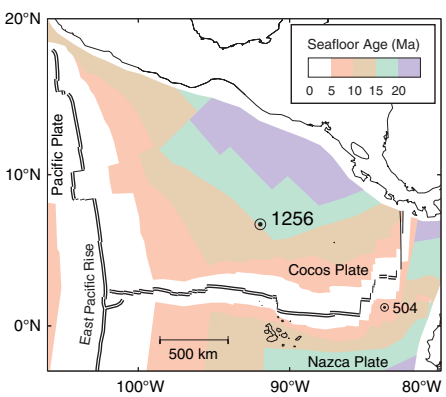


Fig. 2. Age map of the Cocos plate and EPR with isochrons at 5-Ma intervals, converted from magnetic anomaly identifications according to the time scale of Cande and Kent (*49*). The wide spacing of 10- to 20-Ma isochrons to the south reflects the extremely fast (200 to 220 mm/year) full spreading rate. The locations of deep drill holes into the oceanic crust at Sites 1256 and 504 are shown.

[1]Department of Earth Science and Marine Science Institute, University of California, Santa Barbara, CA 93106, USA. [2]National Oceanography Centre, Southampton, University of Southampton, SO14 3ZH, UK. [3]Department of Geological Sciences, University of Michigan, Ann Arbor, MI 48109, USA. [4]Integrated Ocean Drilling Program, Texas A&M University, College Station, TX 77845, USA. [5]Department of Biology and Geosciences, Shizuoka University, Shizuoka 422–8529, Japan. [6]Department of Geology, Niigata University, Niigata 950–2181, Japan. [7]Department of Geology, University of California, Davis, CA 95616, USA. [8]Graduate School of Life and Environmental Sciences, University of Tsukuba, Ibaraki 305–8572, Japan. [9]Department of Geology, University of Leicester, Leicester LE1 7RH, UK. [10]Laboratoire de Géophysique et d'Hydrodynamique en Forage, Institut des Sciences de la Terre, de l'Environnement et de l'Espace de Montpellier, 34095 Montpellier, France. [11]Laboratoire de Géologie, École Normale Supérieure, 75231 Paris, France. [12]College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, OR 97331–5503, USA. [13]Institut Universitaire Européen de la Mer (IUEM), CNRS/UBO, F-29280, Plouzané, France. [14]Dipartimento per lo Studio del Territorio e delle sue Risorse, Università di Genova, 16132 Genova, Italy. [15]Department of Earth Sciences, Florida International University, Miami, FL 33199, USA. [16]Department of Earth Sciences, Università di Milano, 20133 Milano, Italy. [17]Department of Geosciences, University of Houston, Houston, TX 77204, USA. [18]Leibniz Institute for Marine Sciences Kiel, Instituts für Meereskunde–Forschungszentrums für Marine Geowissenschaften, 24148 Kiel, Germany. [19]Maritime Studies Program, Williams College and Mystic Seaport, Mystic, CT 06355, USA. [20]Division of Earth and Ocean Sciences, Duke University, Durham, NC 27708, USA. [21]Hawaii Institute of Geophysics and Planetology, [22]Department of Geology and Geophysics/School of Ocean and Earth Science and Technology, University of Hawaii at Manoa, Honolulu, HI 96822, USA. [23]Graduate School of Environmental Studies, Tohoku University, Sendai 980–8579, Japan. [24]Department of Geology, University of St. Thomas, St. Paul, MN 55105, USA. [25]Department of Geological Sciences, Florida State University, Tallahassee, FL 32306–4100, USA. [26]Bundesanstalt für Geowissenschaften und Rohstoffe, 30655 Hannover, Germany. [27]Department of Geology and Geophysics, University of Utah, Salt Lake City, UT 84112, USA. [28]Institut für Mineralogie, Universität Hannover, 30167 Hannover, Germany. [29]Laboratoire de Pétrologie Magmatique, Université Paul Cézanne Aix-Marseille III, 13397 Marseille 20, France. [30]Department of Geoscience, University of Las Vegas, Las Vegas, NV 89154–4010, USA. [31]Department of Earth Sciences, University of Cambridge, Cambridge CB2 3EQ, UK. [32]School of Earth Sciences, University of Leeds, West Yorkshire LS2 9JT, UK. [33]Polar Research Institute, Incheon 406–840, Korea. [34]Department of Earth and Planetary Science, University of Tokyo, Tokyo 113–0033, Japan. [35]Department of Geology and Paleontology, The National Science Museum, Tokyo 169–0073, Japan. [36]Department of Earth and Planetary Sciences, Macquarie University, Sydney 2109, Australia. [37]Institut für Mineralogie und Geochemie, Universität Karlsruhe, 76131 Karlsruhe, Germany. [38]Department of Geology and Geophysics, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA. [39]Department of Earth and Environmental Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180–3590, USA. [40]Department of Oceanography, Texas A&M University, College Station, TX 77843–3146, USA. [41]Department of Earth and Planetary Sciences, Hokkaido University, Hokkaido 060–0810, Japan. [42]Department of Earth Sciences, Boston University, Boston, MA 02215, USA.

*Ocean Drilling Program (ODP) Leg 206 Shipboard Scientific Party.

†Integrated Ocean Drilling Program (IODP) Expedition 309 Shipboard Scientific Party.

‡IODP Expedition 312 Shipboard Scientific Party.

§To whom correspondence should be addressed. E-mail: dwilson@geol.ucsb.edu

ward from the lowermost dikes into the uppermost gabbro at Hole 1256D. This results from the contact metamorphism of the granoblastic dikes and the strong hydrothermal alteration of the uppermost gabbros (Fig. 3). Porosity and velocity then increase downhole in the gabbro but are still <6.5 km/s.

Flows and dikes from Hole 1256D show a wide range of magmatic fractionation, from fairly primitive to evolved (Figs. 3 and 5). Shallower than 600 mbsf, magma compositions are bimodal, with relatively evolved thick flows and more primitive thin flows. Primitive and evolved compositions are closely juxtaposed within the dikes, as would be expected for vertically intruded magmas. For most major elements and many trace elements, the range of concentrations in flows and dikes is similar to that observed for the northern East Pacific Rise (EPR) (Fig. 5). A few incompatible elements, including Na and Zr, have lower concentrations than observed for modern EPR lavas, but the substantial overlap of compositions indicates similar processes operated at the superfast-spreading ridge that formed Site 1256 and the modern EPR.

The gabbro compositions span a range similar to the flows and dikes but are on average more primitive. Although less fractionated, the average gabbro composition is evolved relative to candidates for primary magma in equilibrium with mantle olivine. Therefore, the residue removed from primary magma to produce the observed gabbro and basalt compositions must be deeper than the uppermost gabbros penetrated in Hole 1256D.

Discussion. Marine seismologists have long been subdividing the ocean crust into seismic layers: Layer 1 has low velocity and is agreed to be sediments; layer 2 has low velocity and high velocity gradient; layer 3 has high velocity (generally at least 6.7 km/s) and low gradient. There is a widespread perception that layer 3 is equivalent to gabbro, even though Hole 504B penetrated layer 3 but not gabbro (*4*, *18*). From regional seismic refraction data the transition from seismic layer 2 to layer 3 at Site 1256 occurs between 1450 to 1750 mbsf (1200 to 1500 msb) (*25*) (Fig. 3). Shipboard determinations of seismic velocities of discrete samples are in close agreement with in situ measurements by wireline tools, and the gabbro velocities are <6.5 km/s. Downhole velocity measurements end at the top of gabbro, but we interpret the gabbro intervals as within layer 2 because a smoothed extrapolation of the downhole velocities will either have velocities <6.5 km/s, still characteristic of layer 2, or will have an exceptionally high gradient to higher velocities, also characteristic of layer 2. Encountering gabbro at a depth clearly within layer 2 reinforces previous suggestions that factors including porosity and alteration are more important than rock type or grain size on controlling the location of the boundary between layers 2 and 3. The position of the dike-gabbro boundary, therefore, has little control over the seismic velocity structure of the crust (*4*, *18*).

Relative to other well-studied upper ocean crust sections (*31*), Site 1256 shows a thick lava sequence and a thin dike sequence. Steady-state thermal models require that the conductive lid separating magma from rapidly circulating seawater thins as spreading rate increases, indicating that the thin dike sequence is a direct consequence of the high spreading rate. A thick flow sequence with many massive individual flows and few pillow lavas is a reasonable consequence of short vertical transport distance from the magma chamber and similar to observations from the middle of segments on the fastest spreading ridges in the modern ocean (*32*). This is in direct contrast to spreading models developed from observations of tectonically disrupted

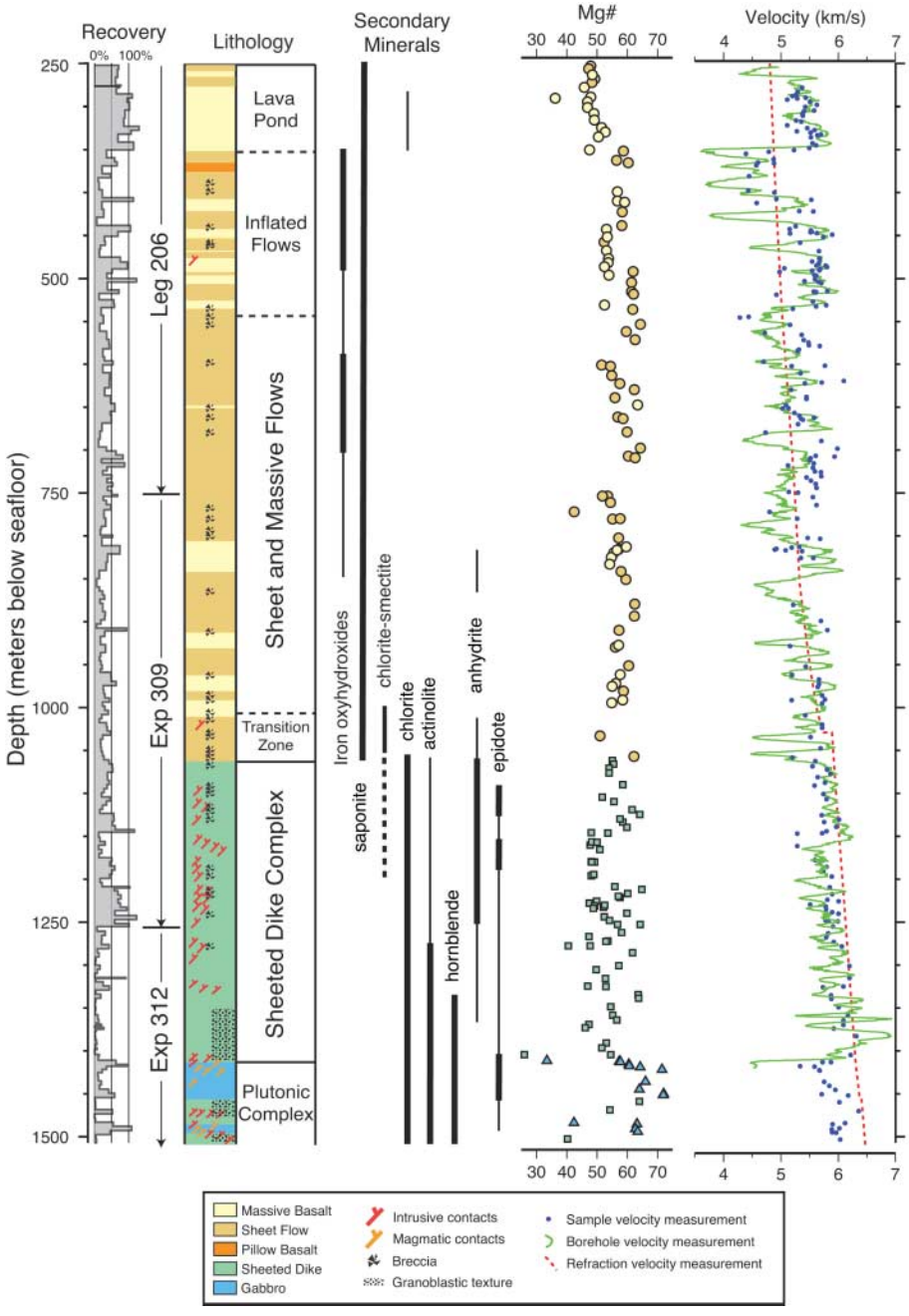


Fig. 3. Summary lithostratigraphic column of the basement drilled to date at Site 1256 showing recovery, major lithologies, downcore index alteration mineral distribution (thick lines show abundant distribution; thin lines show rare distribution), downcore distribution of Mg number [where Mg number = 100 × Mg/(Mg + 0.9 × Fe) atomic ratio; symbols as in Fig. 5], and seismic velocity measured on discrete samples, by wireline tools, and by seismic refraction (*25*).

fast-spread crust exposed in Hess Deep (*33*), which suggest that regions of high magma supply should have thin lavas and thick dikes. Similarly, there is little evidence for tilting (at most a few degrees) in Hole 1256D and no evidence for substantial faulting. In contrast, the upper crust exposed at Hess Deep shows substantial faulting and rotations within the dike complex (*33*). The ponded flow at Site 1256 indicates that faults of ~50 to 100 m must exist in superfast-spread crust to provide the necessary relief for ponding of the flow, but faulting and rotations in the dike section must be less common than in crust formed at fast spreading rates, if observations from Hess Deep are widely applicable.

The simplest model for mid-ocean ridge magma plumbing is that the melt lens imaged by MCS is the magma chamber in which crystal-rich residues are separated from the evolved lavas that reach the seafloor. The upper gabbro, when partially molten, would have had depth and impedance properties consistent with geophysically imaged melt lenses (*14*) based on trends established for active ridges (Fig. 1). If the upper gabbro intruded on axis and extended roughly horizontally for at least hundreds of meters, its MCS image would look exactly like modern melt lens reflectors, yet it could not have been the site of fractionation. Its chilled margin against the underlying dike screen precludes segregating a crystal residue that subsides to form the lower crust as in the gabbro glacier model, and its fractionated composition requires that crystals have been segregated elsewhere. This implies that sills or other bodies containing cumulate materials must exist deeper in the crust and/or below the boundary between crust mantle, consistent with recent models based on lower crustal sections of ophiolites (*5*, *17*) and some marine geophysical experiments (*34*–*37*). However, the gabbro glacier mode of accretion cannot yet be rejected because fractionated gabbros in the dike-gabbro transition are not unexpected, and the predicted region of cumulate rocks could still exist just below the present maximum depth of Hole 1256D.

The ~800-m-thick lava sequence is much less hydrothermally altered than other basement sites [e.g., Sites 417 and 418 and Holes 504B and 896A (*25*, *38*)], and the systematic change with depth from oxidizing to reducing seawater alteration in the upper lavas found elsewhere does not occur. Instead, oxidizing alteration occurs irregularly, associated with steeply dipping vein networks, indicating a structural control of alteration rather than simply decreasing seawater influence downward. The secondary mineralogy of the rocks indicates a stepwise increase in alteration temperatures downhole from ~100°C in the lavas to ~250°C in the uppermost dikes. Aside from the granoblastic

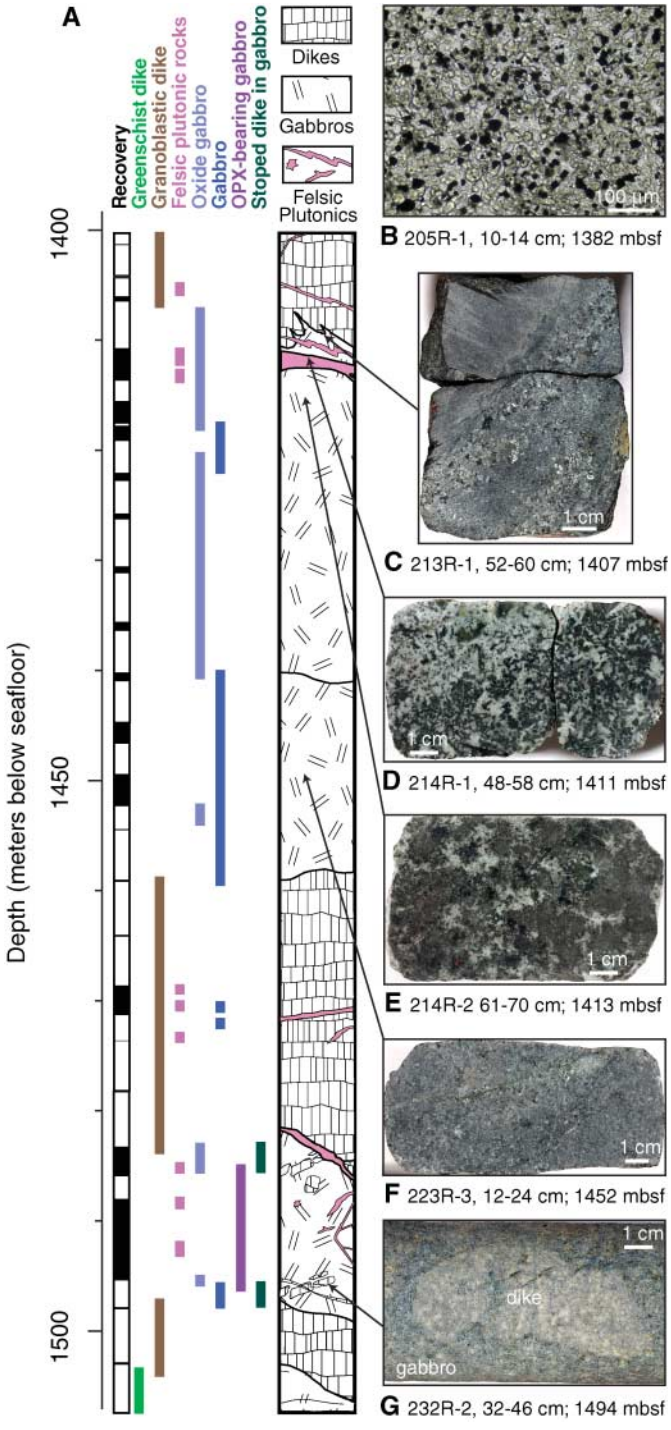


Fig. 4. (**A**) Schematic lithostratigraphic section of the Plutonic Complex from the lower portion of Hole 1256D with representative photographs of key samples. The distribution of rock types is expanded proportionately in zones of incomplete recovery. Felsic plutonic rocks include quartz-rich oxide diorite and trondhjemite. (**B**) Photomicrograph of a dike completely recrystallized to a granoblastic association of equant secondary plagioclase, clinopyroxene, magnetite, and ilmenite. Some granoblastic dikes have minor orthopyroxene. (**C**) The dike-gabbro boundary: Medium-grained oxide gabbro is intruded into granoblastically recrystallized dike along an irregular moderately dipping contact. The gabbro is strongly hydrothermally altered. (**D**) Quartz-rich oxide diorite strongly altered to actinolitic hornblende, secondary plagioclase, epidote, and chlorite. Epidote occurs in ~5-mm clots in the finer grained leucocratic portions of the rock. (**E**) Disseminated oxide gabbro with patchy texture and centimeter-scale dark ophitically intergrown clinopyroxene and plagioclase patches separated by irregular, more highly altered leucocratic zones. (**F**) Medium-grained strongly hydrothermally altered gabbro. The sample is cut by several chlorite + actinolite veins with light gray halos. Plagioclase is replaced by secondary plagioclase and clinopyroxene by amphibole. (**G**) Clast of partially resorbed dike within gabbro.

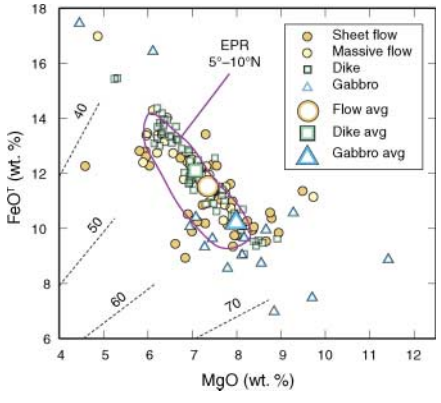


Fig. 5. FeO^T (total Fe expressed as FeO) versus MgO for the basement at Site 1256, compared with analyses of northern EPR (outline) (*50*). Dashed lines show constant Mg number. Possible primary mantle melt compositions should have Mg number of 70 to 78 and MgO of 9 to 14 weight %. All flows and dikes and most gabbros are too evolved to be candidates for primary magmas.

contact metamorphic assemblages in the basal dikes, hydrothermal mineralogy and inferred alteration temperatures of the lower dikes in Hole 1256D are generally similar to those in the lower dikes of Hole 504B (up to ~400°C). However, the dike section at Site 1256 is much thinner than the section at Site 504 (~350 compared with >1000 m), which indicates a much steeper hydrothermal temperature gradient at Site 1256 (~0.5°C/m compared with 0.16°C/m in 504B).

Forming the lower oceanic crust through the crystallization and subsidence of a high-level melt lens (*6*, *7*) is the most efficient geometry for hydrothermal cooling of the crust, as the latent and sensible heat can be readily advected by shallowly circulating hydrothermal fluids. However, models have predicted larger hydrothermal fluid fluxes and more intense alteration than those that have been documented from ocean crustal sections to date (*39*–*41*). Epidosites—equigranular epidote-quartz-titanite rocks that delineate zones of intense hydrothermal leaching and channel-ways of upwelling black smoker-type fluids—are common in ophiolites around the dike-gabbro boundary (*42*–*44*) and are recorded in the oceans from fore-arc crust (*45*). Although epidote is a common replacement mineral within and below the transition zone in Hole 1256D (Fig. 3) and rare epidote-rich alteration patches are present, epidosites were not recovered. Because of the retrograde solubility of calcium sulfate, anhydrite precipitation must play a critical but to date poorly understood role in axial hydrothermal circulation (*46*, *47*). Anhydrite is more abundant in the lava-dike transition and in the upper dikes in Hole 1256D (Fig. 3) compared with Hole 504B (*47*), but still present in much lower quantities than predicted by numerical models of hydrothermal circulation resulting from the crystallization of the crust in an axial melt lens (*46*).

The question of whether the lower crust solidifies by shallow crystallization and subsidence (*6*) or in situ crystallization cooled by deep hydrothermal circulation (*9*) remains unresolved. The former model requires more intense hydrothermal circulation above a shallow magma chamber to remove the latent heat from crystallizing the lower crust; the latter requires less shallow circulation but more total circulation to remove both latent and sensible heat from the lower crust. The high thermal gradient inferred for the dike layer appears adequate to conduct latent heat of crystallization from a shallow magma chamber, but evidence such as epidosites for the large volume of water required to react with hot rock to sustain this thermal gradient is lacking. Also, the inferred thermal gradient may not reflect steady-state conditions, with high temperatures at depth possibly transitory as the shallow gabbro crystallized. However, evidence for the large volumes of water that must pass through the upper crust to reach and cool the lower crust is also lacking. Retrograde metamorphism under strongly hydrous conditions would be expected below about 1000 msb as large volumes of water quenched the upper crust before circulating at depth.

The penetration of the dike-gabbro boundary in intact oceanic crust is a major milestone that has taken more than four decades of scientific ocean drilling that began with Operation MoHole (*48*). Our success shows that deep drilling in young ocean crust formed at fast spreading rates is feasible. Such deep drill holes are key to testing models of crustal accretion and calibrating observations from marine geophysics and ophiolites. The shallow depth to gabbros at Site 1256 was predicted from seismic studies of modern active ridges by extrapolation to a spreading rate substantially faster than that occurring on the planet today. Further deepening of Hole 1256D will challenge current ideas on the formation and cooling of the lower crust, for example, testing the hypothesis that seismic layer 3 at this site consists of relatively impermeable fresh gabbro that cooled by conduction.

References and Notes

1. J. M. Sinton, R. S. Detrick, *J. Geophys. Res.* **97**, 197 (1992).
2. D. R. Toomey, G. M. Purdy, S. C. Solomon, W. S. D. Wilcock, *Nature* **347**, 639 (1990).
3. E. E. Vera *et al.*, *J. Geophys. Res.* **95**, 15529 (1990).
4. R. S. Detrick, J. A. Collins, R. A. Stephen, S. A. Swift, *Nature* **370**, 288 (1994).
5. C. J. MacLeod, G. Yoauancq, *Earth Planet. Sci. Lett.* **176**, 357 (2000).
6. T. J. Henstock, A. W. Woods, R. S. White, *J. Geophys. Res.* **98**, 4143 (1993).
7. J. Phipps Morgan, Y. J. Chen, *J. Geophys. Res.* **98**, 6283 (1993).
8. J. E. Quick, R. P. Denlinger, *J. Geophys. Res.* **98**, 14015 (1993).
9. J. Maclennan, T. Hulme, S. C. Singh, *Geology* **33**, 357 (2005).
10. R. S. Detrick *et al.*, *Nature* **326**, 35 (1987).
11. R. S. Detrick *et al.*, *Science* **259**, 499 (1993).
12. G. M. Kent, A. J. Harding, J. A. Orcutt, *Nature* **344**, 650 (1990).
13. J. C. Mutter *et al.*, *Nature* **336**, 156 (1988).
14. S. C. Singh, G. M. Kent, J. S. Collier, A. J. Harding, J. A. Orcutt, *Nature* **394**, 874 (1998).
15. S. Carbotte, C. Mutter, J. C. Mutter, G. Ponce-Correa, *Geology* **26**, 455 (1997).
16. G. M. Purdy, L. S. L. Kong, G. L. Christeson, S. C. Solomon, *Nature* **355**, 815 (1992).
17. P. B. Kelemen, K. Koga, N. Shimizu, *Earth Planet. Sci. Lett.* **146**, 475 (1997).
18. J. C. Alt *et al.*, *Proc. ODP Sci. Res.* **148**, 417 (1996).
19. H. J. B. Dick, J. H. Natland, D. J. Miller *et al.*, *Proceedings of the Ocean Drilling Program, Initial Reports* (Ocean Drilling Program, College Station, TX, 1999), vol. 176, [CD-ROM].
20. P. B. Kelemen, E. Kikawa, D. J. Miller *et al.*, *Proceedings of the Ocean Drilling Program Initial Reports* (Ocean Drilling Program, College Station, TX, 2004), vol. 209, [CD-ROM].
21. K. Gillis *et al.*, *Proceedings of the Ocean Drilling Program, Initial Reports* (Ocean Drilling Program, College Station, TX, 1993), vol. 147.
22. Expedition 304 Scientists, *IODP Prel. Rep.* **304**, 10.2204/iodp.pr.304.2005 (2005).
23. Expedition 305 Scientists, *IODP Prel. Rep.* **305**, 10.2204/iodp.pr.305.2005 (2005).
24. D. S. Wilson, *Geophys. Res. Lett.* **23**, 3003 (1996).
25. D. S. Wilson *et al.*, *Proc. ODP Init. Repts.* (Ocean Drilling Program, College Station, TX, 2003), vol. **206** [CD-ROM].
26. Expedition 309 Scientists, *IODP Prel. Rep.* **309**, 10.2204/iodp.pr.309.2005 (2005).
27. Expedition 309 and 312 Scientists, *IODP Prel. Rep.* **312**, 10.2204/iodp.pr.312.2006 (2006).
28. K. C. Macdonald, P. J. Fox, R. T. Alexander, R. Pockalny, P. Gente, *Nature* **380**, 125 (1996).
29. S. Umino, S. Obata, P. W. Lipman, *Geology* **28**, 503 (2000).
30. E. Hallenborg, A. J. Harding, G. M. Kent, D. S. Wilson, *J. Geophys. Res.* **108**, 2532 10.1029/2003JB002400 (2003).
31. J. A. Karson, *Annu. Rev. Earth Planet. Sci.* **30**, 347 (2002).
32. S. M. White, R. M. Haymon, D. J. Fornari, M. R. Perfit, K. C. Macdonald, *J. Geophys. Res.* **107**, 2173 10.1029/2001JB000571 (2002).
33. J. A. Karson *et al.*, *Geochem. Geophys. Geosyst.* **3**, 1002 10.1029/2001GC000155 (2002).
34. W. C. Crawford, S. C. Webb, *Earth Planet. Sci. Lett.* **203**, 117 (2002).
35. R. A. Dunn, D. R. Toomey, R. S. Detrick, W. S. D. Wilcock, *Science* **291**, 1955 (2001).
36. J. Garmany, *Nature* **340**, 628 (1989).
37. M. R. Nedimovic *et al.*, *Nature* **436**, 1149 (2005).
38. J. C. Alt, in *Hydrogeology of the Oceanic Lithosphere*, H. Elderfield, E. Davis, Eds. (Cambridge Univ. Press, New York, 2004), pp. 456–488.
39. J. C. Alt, D. A. H. Teagle, *Geol. Soc. Am. Spec. Pap.* 349 (2000), pp. 273–282.
40. K. M. Gillis, L. A. Coogan, R. Pedersen, *Earth Planet. Sci. Lett.* **232**, 83 (2005).
41. D. A. H. Teagle, M. J. Bickle, J. C. Alt, *Earth Planet. Sci. Lett.* **210**, 81 (2003).
42. P. Nehlig, T. Juteau, V. Bendel, J. Cotten, *J. Geophys. Res.* **99**, 4703 (1994).
43. C. J. Richardson, J. R. Cann, H. G. Richards, J. G. Cowan, *Earth Planet. Sci. Lett.* **84**, 243 (1987).
44. P. Schiffman, B. M. Smith, R. J. Varga, E. M. Moores, *Nature* **325**, 423 (1987).
45. N. R. Banerjee, K. M. Gillis, K. Muehlenbachs, *Geology* **28**, 151 (2000).
46. N. H. Sleep, *J. Geophys. Res.* **96**, 2375 (1991).
47. D. A. H. Teagle, J. C. Alt, A. N. Halliday, *Earth Planet. Sci. Lett.* **155**, 167 (1998).
48. W. Bascom, *A Hole in the Bottom of the Sea* (Doubleday and Co., New York, 1961), pp. 352.
49. S. C. Cande, D. V. Kent, *J. Geophys. Res.* **100**, 6093 (1995).
50. C. H. Langmuir, J. F. Bender, R. Batiza, *Nature* **322**, 422 (1986).

51. This research used samples and data provided by the ODP and the IODP. ODP was sponsored by the NSF and participating countries under management of Joint Oceanographic Institutions (JOI), Inc. IODP is supported by NSF; Japan's Ministry of Education, Culture, Sports, Science, and Technology; the European Consortium for Ocean Research Drilling; and the People's Republic of China, Ministry of Science and Technology. For Leg 206, D.S.W. and D.A.H.T. were co-chief scientists, and G.D.A. was the staff scientist. For Expedition 309, D.A.H.T. and S.U. were co-chief scientists, and N.R.B. was the staff scientist. For Expedition 312, J.C.A. and S.M. were co-chief scientists, and N.R.B. was the staff scientist. We thank *JOIDES Resolution* Captains P. Mowat and A. Simpson, Drilling Superintendents Tim McCown and Wayne Malone, Operations Superintendents R. Grout and K. Grigar, the entire transocean crew, and the Texas A&M University technical staff from ODP Leg 206 and IODP Expeditions 309 and 312 for successful operations.

10 February 2006; accepted 6 April 2006
Published online 20 April 2006;
10.1126/science.1126090
Include this information when citing this paper.

Spin Coupling in Engineered Atomic Structures

Cyrus F. Hirjibehedin,* Christopher P. Lutz, Andreas J. Heinrich

We used a scanning tunneling microscope to probe the interactions between spins in individual atomic-scale magnetic structures. Linear chains of 1 to 10 manganese atoms were assembled one atom at a time on a thin insulating layer, and the spin excitation spectra of these structures were measured with inelastic electron tunneling spectroscopy. We observed excitations of the coupled atomic spins that can change both the total spin and its orientation. Comparison with a model spin-interaction Hamiltonian yielded the collective spin configuration and the strength of the coupling between the atomic spins.

Meticulous control of the coupling between spins is central to the use of magnetic nanostructures in future spintronics (*1*) and quantum information devices (*2*). Such spin coupling has been studied recently in pairs of lithographically fabricated quantum dots (*3*, *4*). At the atomic length scale, a small number of spins can be coupled directly in metal-atom clusters (*5*) or chains (*6*) and through linker groups in molecular magnetic structures (*7*). Spin interactions in these systems typically have been probed using ensemble-averaging techniques, such as susceptibility measurements (*8*, *9*), electron paramagnetic resonance (*10*, *11*), and inelastic neutron scattering (*12*).

Individual atomic-scale magnetic structures can be studied with scanning probe techniques (*13*, *14*). The scanning tunneling microscope (STM) has been used to detect the Kondo interactions between conduction electrons and single atomic spins (*15*, *16*) and to study the properties of individual magnetic islands (*17*, *18*). Recently, inelastic electron tunneling spectroscopy (IETS) with an STM was used to measure the spin excitation spectra of individual magnetic atoms (*19*). Additionally, the STM can be used to manipulate and assemble individual nanostructures (*20*). Such atomic manipulation in situ was used to construct magnetic dimers and trimers, which displayed evidence of coupled-spin behavior (*21*, *22*).

We have combined the imaging and manipulation capabilities of the STM with its ability to measure spin excitation spectra at the atomic scale. We used an STM to build magnetic chains as long as 10 Mn atoms on a thin insulating layer and probed their collective spin excitations with IETS. The spectra depended strongly on both the parity and length of the chains, as well as the binding sites upon which the chains were built. We observed excitations that can change both the collective spin of the coupled system and the orientation of the spin with respect to a magnetic field. A quantitative comparison with the Heisenberg open chain model allowed us to determine the antiferromagnetic (AF) coupling strength, which could be set at either 3 or 6 meV by selecting the precise atomic configuration. In addition, the spin of the Mn atoms in the chains was found to be ½.

Experiments were conducted with ultrahigh-vacuum, low-temperature STMs reaching base temperatures of 5.5 and 0.6 K. In the latter STM, we applied magnetic fields up to $B = 7$ T in the plane of the sample, oriented ~55° from the chains. We measured differential conductances dI/dV using lock-in detection of the tunnel current I by adding a 20 to 50 μV_{rms} modulation at 732 Hz to the sample bias voltage V.

To date, spin excitations have been observed with IETS only when the atomic spins were separated from the metal surface by a thin insulating layer (*19*). Here we used the insulating properties of a single atomic layer of copper nitride (*23*, *24*), which we denote as CuN. A small island of CuN surrounded by bare Cu(100) is shown in Fig. 1A, along with the assigned positions of the surface atoms. The relatively simple structure of these islands offers a well-characterized environment in which to assemble nanostructures.

Prior STM studies using thin insulating layers (*19*, *25*–*27*) did not report the ability to precisely position atoms or molecules. We have developed a vertical transfer technique (*28*, *29*) to manipulate Mn atoms on the CuN surface with atomic precision (*30*). In the atom-by-atom construction of a chain of 10 Mn atoms on a

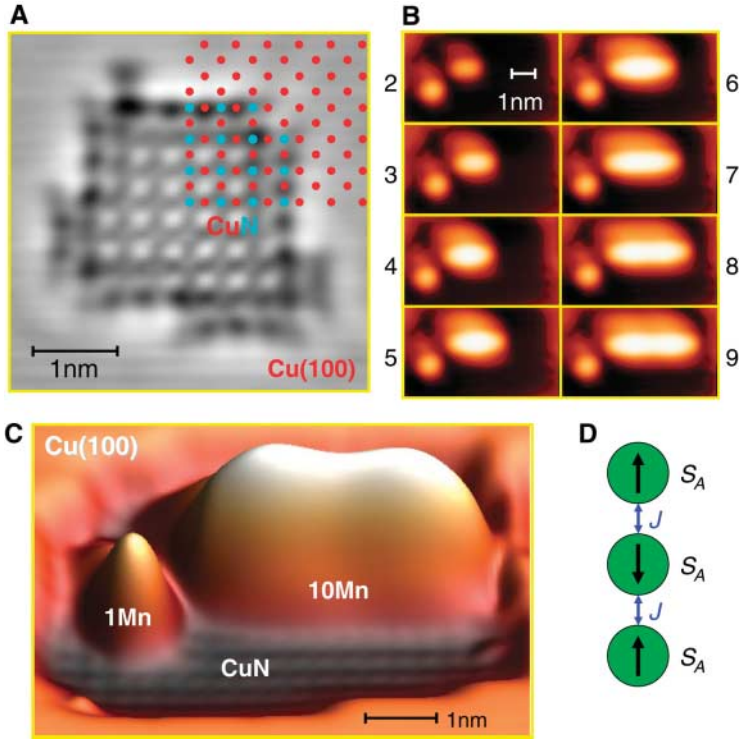


Fig. 1. Mn chains on CuN. (**A**) STM constant-current topograph (10 mV, 1 nA) of a CuN island on Cu(100). Topograph is high-pass (curvature) filtered to enhance contrast, with lattice positions of Cu (red dots) and N (blue dots) atoms overlaid. CuN islands appear as 0.14-nm depressions for $|V| < 0.1$ V. (**B**) STM images of the building of a chain of Mn atoms, lengths 2 to 9, on CuN (10 mV, 0.1 nA). The Mn atom in the lower left of each image appears as a bump 0.34 nm high. The dimer has a smaller peak height of 0.30 nm. Chains of four or more atoms have peak heights of ~0.45 nm. Individual atoms in the chain cannot be resolved. Artifacts seen to the upper left of all structures are characteristic of the atomic arrangement of the tip used for manipulation. (**C**) Perspective rendering of a chain of 10 Mn atoms. The observed double-peak structure suggests the existence of one-dimensional delocalized electronic states as seen in metallic chains constructed on metal surfaces (*38*). (**D**) Schematic of the AF coupling of three atomic spins described by the Heisenberg model in Eq. 1.

IBM Research Division, Almaden Research Center, 650 Harry Road, San Jose, CA 95120, USA.

*To whom correspondence should be addressed. E-mail: hirjibe@us.ibm.com

CuN island (Fig. 1, B and C), the Mn atoms were positioned atop Cu atoms in the CuN so that one N atom lay between each pair of Mn binding sites. As a result, the Mn atoms were spaced 0.36 nm apart.

The dI/dV spectra of Mn chains of 1 to 10 atoms at $B = 0$ (Fig. 2) displayed a striking dependence on the parity of the chain. At voltages below 1 mV, the spectra of odd-length chains, including the single Mn atom, exhibited a narrow dip in conductance that was centered at 0 V. This feature was absent in all even-length chains. At larger energies, both odd and even chains, with the exception of single Mn atoms, showed large steps in conductance at voltages that were symmetric with respect to zero. These steps were more pronounced in the long even-length chains: The change in conductance across these steps was almost a full order of magnitude. The energies of these steps were largest for the dimer and trimer, and they decreased with chain length for chains of the same parity. The conductance steps occurred at the same energy when the spectra were measured at different locations along the length of the chains, which suggests that these steps are properties of the entire chain rather than of any one atom or part of the chain.

We used IETS (*31*, *32*) to interpret the steps in the dI/dV spectra of the chains. When a system with a mode at energy E is placed within a tunnel junction, electrons can excite this mode during the tunneling process only if the magnitude of the bias voltage V exceeds E/e (where $-e$ is the charge on the electron). This additional tunneling channel generally results in steps in conductance at $V = \pm E/e$.

To understand the IETS steps in Fig. 2, we focus first on the case of a single Mn atom. The dI/dV spectra on a single Mn atom on CuN as a function of magnetic field (Fig. 3) all exhibit a dip centered at $V = 0$ that we interpret as symmetric IETS steps in conductance, similar to the spin-flip excitations observed for Mn on aluminum oxide (*19*). A spin-flip excitation changes the magnetic quantum number m but leaves the total spin S unchanged. The steps in conductance shift to higher voltage with increasing magnetic field as described by $eV = \pm g\mu_B B$ with $g \sim 2$, where g is the Landé g value and μ_B is the Bohr magneton. The odd-length chains all show similar behavior around $V = 0$, as illustrated for the trimer in Fig. 3B. The existence of spin-flip excitations requires $S > 0$ for the ground state of these chains. In contrast to previous work (*19*), the spin-flip excitations seen in Fig. 3A deviate at low magnetic fields from the expected energy of $g\mu_B B$: As seen in Fig. 2, a spin-flip excitation is present even at $B = 0$ in the odd-length chains. This zero-field splitting grows in energy with increasing chain length and is likely caused by magnetocrystalline anisotropy arising from the loss of rotational symmetry at the surface (*33*, *34*).

None of the even-length chains exhibited a spin-flip excitation: the spectra are flat near $V = 0$, both at $B = 0$ and at fields up to $B = 7$ T (Fig. 4A). The absence of the spin-flip excitation on the even-length chains is consistent with a ground-state spin $S = 0$. The alternation of the ground-state spin between zero for even-length chains and nonzero values for odd-length chains indicates AF coupling between the atomic spins.

For the even-length chains in Fig. 2, we interpret the lowest energy conductance steps as excitations that change the total spin of the chain. The dimer's conductance spectra between 4 and 8 mV are shown as a function of magnetic field in Fig. 4A. The single step at 6 mV at zero field is split into three distinct steps at $B = 7$ T; we interpret these steps as three IETS excitations. The step energies, as determined by IETS fitting, are shown in Fig. 4B. The three excitations shift as $\Delta m g\mu_B B$ for $\Delta m = -1, 0, 1$ and $g = 2.1 \pm 0.1$. The emergence of three distinct excitations at large B that are essentially degenerate at $B = 0$ suggests that the first excited state for the even-length chains is a triplet state ($S = 1$). The IETS steps thus correspond to spin-changing transitions from the ground state (with $S = 0$ and $m = 0$) to an excited state (with $S = 1$ and $m = -1, 0, 1$) (Fig. 4C). We interpret the different zero-field energy of the $m = \pm 1$ and $m = 0$ excited states as another manifestation of magnetocrystalline anisotropy. A similar singlet-triplet excitation

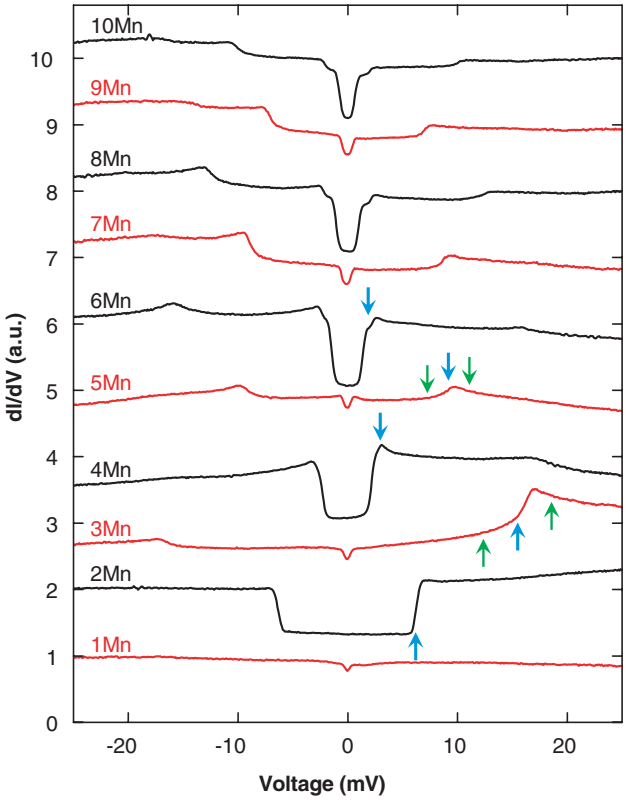


Fig. 2. Conductance spectra of Mn chains on CuN. Spectra were taken with the tip positioned above the center of chains of Mn atoms of lengths 1 to 10 at $T = 0.6$ K and $B = 0$ T. Spectra were acquired at a nominal junction impedance of 20 megohms (20 mV, 1 nA) and were not sensitive to junction impedance. Successive spectra are vertically offset by one unit for clarity. Odd spectra are in red and even spectra are in black to emphasize the parity dependence. Blue arrows indicate the lowest energy spin-changing excitation obtained from the Heisenberg model described in Eq. 1 with $J = 6.2$ meV and $S_A = ½$. For comparison, green arrows indicate the same excitation with $S_A = 2$ (lower voltage) and $S_A = 3$ (higher voltage) on the trimer and pentamer; the same excitation with $S_A = 2$ and $S_A = 3$ is not distinguishable from $S_A = ½$ on the dimer, tetramer, and hexamer chains.

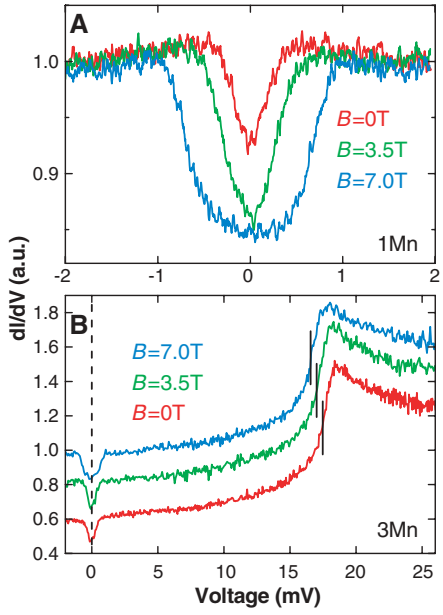


Fig. 3. Conductance spectra of Mn atom and trimer on CuN. (**A**) Spectra taken on a single Mn atom on CuN at $T = 0.6$ K at $B = 0$, 3.5, and 7.0 T. All spectra were acquired at a nominal junction impedance of 20 megohms (8 mV, 0.4 nA). (**B**) Spectra taken on a Mn trimer on CuN at $T = 0.6$ K at $B = 0$, 3.5, and 7.0 T. All spectra were acquired at a nominal junction impedance of 20 megohms (20 mV, 1 nA), and successive spectra are vertically offset by 0.2 units for clarity. Vertical lines highlight the shift of the high-energy step to lower energy with increasing magnetic field.

also occurs for the other even-length chains. As seen in Fig. 2, the zero-field splitting becomes more pronounced in the longer even-length chains, where two distinct steps are well resolved even at $B = 0$.

We analyze our spectroscopic results by using the Heisenberg model for an open chain of N coupled spins. In its simplest form, this model neglects any zero-field splitting and includes only identical nearest neighbor exchange interactions (coupling strength J) between the spins (Fig. 1D). We assume that all sites have the same spin S_A. The Hamiltonian is

$$H_N = J\sum_{i=1}^{N-1} \mathbf{S}_i \cdot \mathbf{S}_i + 1 \qquad (1)$$

where $\mathbf{S}_i$ is the spin operator for the ith site along the chain. The eigenvalue of $\mathbf{S}_i^2$ is thus $S_A(S_A + 1)$ for every i.

For a dimer, the ground state of the AF ($J > 0$) Heisenberg chain is a singlet ($S = 0$) and the first excited state is a triplet ($S = 1$) with energy separation J. This result is independent of the spin S_A of the constituent atoms. The energy spacing between the ground state and the first excited state provides a direct measure of the coupling strength J. For the dimer shown in Fig. 2, we determine $J = 6.2$ meV. We find that the value of J varies by ±5% for dimers in different locations on several CuN islands, as can be seen for the dimer in Fig. 4, which has $J = 5.9$ meV, and in fig. S1A, which has $J = 6.4$ meV.

In contrast to the dimer, the ground state of the Heisenberg AF trimer has total spin $S = S_A$. Generally, the first excited state has total spin $S_A - 1$ and is higher in energy by $S_A J$; the exception is the case of $S = ½$, where both the ground and first excited states have total spin ½. Figure 2 shows the energy of the first excited state for the trimer for a few possible values of S_A, where the value of J used was that determined from the dimer. The conductance step at ~16 mV, which we now assign to the lowest spin-changing transition, is best matched with the result for $S_A = ½$. The presence of a broad peak in the trimer spectrum, centered near the IETS step voltage, lies outside the IETS framework and makes a quantitative fit to the step position difficult. We note that $S_A = ½$ is identical to the spin of a free Mn atom.

From the excitation spectra of the dimer and trimer, we determined both free parameters J and S_A of the Heisenberg Hamiltonian in Eq. 1. Using these two values, we then calculated the excitation spectra for chains up to $N = 6$ (*35*). For odd-length chains, we found that the ground and first excited states have total spin ½ and ³⁄₂, respectively. In contrast, even-length chains have ground and first excited state spins of 0 and 1, respectively. This result is consistent with the observed parity dependence of the spin-flip feature and with the triplet splitting of the spin-changing excitation of even-length chains. The calculated energy of the first excited state, as shown in Fig. 2, is in very good agreement with the observed spectral features on the longer chains, considering the simplicity of the model.

To further demonstrate the potential to engineer the coupling between atomic spins, we have also built dimers with the same Mn spacing but atop N atoms in the CuN layer. Conductance spectra of such a dimer also show a single IETS step, now at ~3 mV, that splits into three distinct steps in a magnetic field. As seen in fig. S1, the coupling strength in these dimers is half as large as for dimers built atop Cu sites.

We note that only a single spin-flip and a single spin-changing transition were observed on odd-length chains. For example, in the trimer spectra shown in Fig. 3B, only a single IETS step is observed for each type of transition; at large B, the higher energy IETS step does not split into multiple steps but simply shifts to lower energy. Because $S = ½$ for the ground state and ³⁄₂ for the first excited state, we would expect to observe multiple spin-flip and spin-changing transitions in a magnetic field. It is thus likely that some transitions are prohibited by IETS selection rules. By using the spin states determined from the Heisenberg model, we find IETS steps only when $\Delta S = 0, \pm 1$ and $\Delta m = 0, \pm 1$. A more detailed understanding of the inelastic tunneling process may shed light on the origins of these selection rules.

One aspect of the spectra of the Mn chains that is not currently understood is the clear asymmetry with respect to voltage polarity. In the trimer spectrum in Fig. 2, the conductance steps occur at both positive and negative voltages but the step heights are markedly different. This asymmetry enhances features at positive voltages for small chains and then shifts to enhance the negative voltage features on longer chains.

Many of the zero-field 0.6 K results presented here were also observed with our 5.5 K STM, with the expected thermal broadening of the spectra, which implies that useful IETS studies of exchange-coupled magnetic nanostructures may be performed without the requirement of subkelvin instruments. A comparison with ab initio calculations may help to determine the nature of the observed exchange coupling, which could be the result of superexchange through molecular orbitals or mediated by the conduction electrons of the underlying metal substrate.

We have shown that the combination of in situ imaging, atom manipulation, and spin-excitation spectroscopy can be used to assemble and probe magnetic nanostructures. Such structures may serve as new model systems for the exploration of low-dimensional magnetism (*36*), and may be used to realize spin-based computation schemes (*37*) and bistable systems for information storage at the atomic scale.

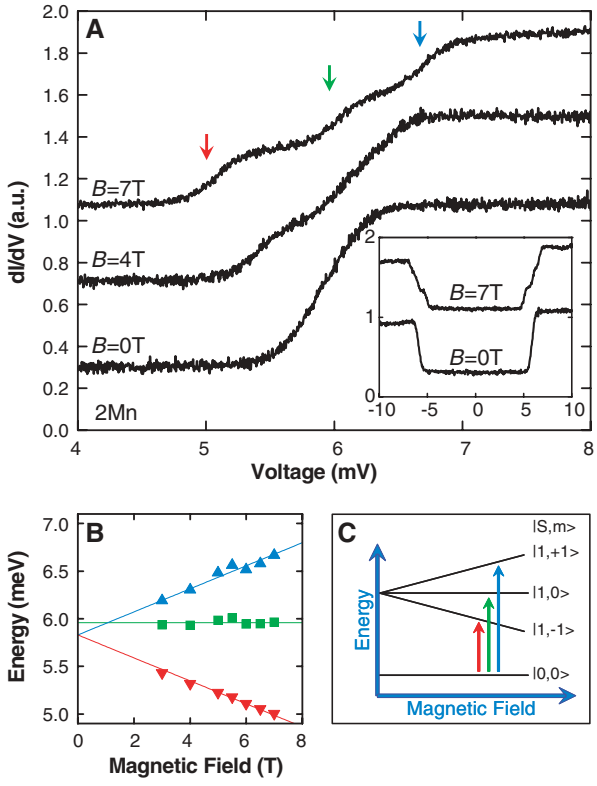


Fig. 4. Conductance spectra of Mn dimer on CuN. (**A**) Spectra taken on a Mn dimer on CuN at $T = 0.6$ K at $B = 0$, 4, and 7 T. All spectra were acquired at a nominal junction impedance of 20 megohms (8 mV, 0.4 nA), and successive spectra are vertically offset by 0.4 units for clarity. Arrows indicate the positions of the low (red), central (green), and high (blue) energy steps from fits of the 7 T spectrum with IETS line shapes (*31*). Inset shows the same spectra, taken at a nominal junction impedance of 50 megohms (20 mV, 0.4 nA), over a larger voltage range. (**B**) Step energies from IETS fits for low (red triangles), central (green squares), and high (blue triangles) energy steps in the dimer spectra as a function of magnetic field. Best fits were obtained with temperatures of ~1.0 K. The uncertainty of the points is typically smaller than the size of the symbols. The solid lines are fits of the step energies to $E_{|\Delta m|} + \Delta m g \mu_B B$, with $g = 2.1 \pm 0.1$, $E_0 = 5.96 \pm 0.05$ meV for $\Delta m = 0$ (green line), and $E_1 = 5.83 \pm 0.05$ meV for $\Delta m = \pm 1$ (blue and red lines). (**C**) Schematic of the energy-level spacing of singlet and triplet states in a magnetic field, labeled by total spin S and magnetic quantum number m. Transitions from the $|0,0\rangle$ singlet state to the $|1,-1\rangle$, $|1,0\rangle$, and $|1,+1\rangle$ triplet states are labeled by red, green, and blue arrows, respectively.

References and Notes

1. S. A. Wolf *et al.*, *Science* **294**, 1488 (2001).
2. M. N. Leuenberger, D. Loss, *Nature* **410**, 789 (2001).
3. F. H. L. Koppens *et al.*, *Science* **309**, 1346 (2005).
4. J. R. Petta *et al.*, *Science* **309**, 2180 (2005).
5. X. Xu, S. Yin, R. Moro, W. A. de Heer, *Phys. Rev. Lett.* **95**, 237209 (2005).
6. P. Gambardella *et al.*, *Nature* **416**, 301 (2002).
7. O. Kahn, *Molecular Magnetism* (Wiley-VCH, New York, 1993).
8. J. R. Friedman, M. P. Sarachik, J. Tejada, R. Ziolo, *Phys. Rev. Lett.* **76**, 3830 (1996).
9. L. Thomas *et al.*, *Nature* **383**, 145 (1996).
10. A. Abragan, B. Bleaney, *Electron Paramagnetic Resonance of Transition Ions* (Clarendon, Oxford, 1970).
11. S. Hill, R. S. Edwards, N. Aliaga-Alcalde, G. Christou, *Science* **302**, 1015 (2003).
12. R. Caciuffo *et al.*, *Phys. Rev. Lett.* **81**, 4744 (1998).
13. D. Rugar, R. Budakian, H. J. Mamin, B. W. Chui, *Nature* **430**, 329 (2004).
14. A. Zhao *et al.*, *Science* **309**, 1542 (2005).
15. V. Madhavan, W. Chen, T. Jamneala, M. F. Crommie, N. S. Wingreen, *Science* **280**, 567 (1998).
16. J. Li, W.-D. Schneider, R. Berndt, B. Delley, *Phys. Rev. Lett.* **80**, 2893 (1998).
17. M. Bode, O. Pietzsch, A. Kubetzka, R. Wiesendanger, *Phys. Rev. Lett.* **92**, 067201 (2004).
18. A. Yamasaki *et al.*, *Phys. Rev. Lett.* **91**, 127201 (2003).
19. A. J. Heinrich, J. A. Gupta, C. P. Lutz, D. M. Eigler, *Science* **306**, 466 (2004).
20. D. M. Eigler, E. K. Schweizer, *Nature* **344**, 524 (1990).
21. T. Jamneala, V. Madhavan, M. F. Crommie, *Phys. Rev. Lett.* **87**, 256804 (2001).
22. H. J. Lee, W. Ho, M. Persson, *Phys. Rev. Lett.* **92**, 186802 (2004).
23. F. M. Leibsle, C. F. J. Flipse, A. W. Robinson, *Phys. Rev. B* **47**, 15865 (1993).
24. Small islands of copper nitride were formed on Cu(100) by nitrogen implantation using a 1-keV beam of N_2 ions followed by annealing to 400°C for 1 min. The sample was then transferred into the precooled STM, and Mn atoms were subsequently evaporated onto the surface.
25. X. H. Qiu, G. V. Nazin, W. Ho, *Science* **299**, 542 (2003).
26. E. Ćavar *et al.*, *Phys. Rev. Lett.* **95**, 196102 (2005).
27. J. Repp, G. Meyer, S. M. Stojković, A. Gourdon, C. Joachim, *Phys. Rev. Lett.* **94**, 026803 (2005).
28. D. M. Eigler, C. P. Lutz, W. E. Rudge, *Nature* **352**, 600 (1991).
29. L. Bartels, G. Meyer, K.-H. Rieder, *Phys. Rev. Lett.* **79**, 697 (1997).
30. Mn was transferred to the tip from either the CuN or bare Cu surface by moving the tip into point contact with the atom and then withdrawing the tip while applying +2 V. The atom was then transferred to the CuN by repeating this procedure with –0.5 V. Mn atoms can bind to the CuN atop either a Cu or N atom, with the Cu site being more stable.
31. J. Lambe, R. C. Jaklevic, *Phys. Rev.* **165**, 821 (1968).
32. B. C. Stipe, M. A. Rezaei, W. Ho, *Science* **280**, 1732 (1998).
33. P. Gambardella *et al.*, *Science* **300**, 1130 (2003).
34. The conductance steps at low fields on Mn atoms and trimers are too narrow to resolve the flat bottom that clearly differentiates an IETS line shape from the Fano line shape that characterizes the Kondo effect (*15*, *16*). Because this flat bottom is seen on longer chains, we consider the Kondo interpretation to be unlikely for the shorter chains as well.
35. Because H_N is represented as a $(2S_A + 1)^N \times (2S_A + 1)^N$ matrix, numerical diagonalization becomes computationally expensive for larger S_A and N. We therefore limited our calculations to $N \leq 6$.
36. L. J. de Jongh, A. R. Miedema, *Adv. Phys.* **23**, 1 (1974).
37. A. Imre *et al.*, *Science* **311**, 205 (2006).
38. N. Nilius, T. M. Wallis, W. Ho, *Science* **297**, 1853 (2002).
39. STM topographic images were processed using WSxM (www.nanotec.es). We thank D. M. Eigler for mentoring and stimulating discussions; B. J. Melior for expert technical assistance; and C.-Y. Lin and B. A. Jones for stimulating discussions.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1125398/DC1
Fig. S1
References and Notes

25 January 2006; accepted 13 March 2006
Published online 30 March 2006;
10.1126/science.1125398
Include this information when citing this paper.

Desorption of H from Si(111) by Resonant Excitation of the Si-H Vibrational Stretch Mode

Zhiheng Liu,[1,2] L. C. Feldman,[2,3] N. H. Tolk,[2] Zhenyu Zhang,[3,4] P. I. Cohen[1]*

Past efforts to achieve selective bond scission by vibrational excitation have been thwarted by energy thermalization. Here we report resonant photodesorption of hydrogen from a Si(111) surface using tunable infrared radiation. The wavelength dependence of the desorption yield peaks at 0.26 electron volt: the energy of the Si-H vibrational stretch mode. The desorption yield is quadratic in the infrared intensity. A strong H/D isotope effect rules out thermal desorption mechanisms, and electronic effects are not applicable in this low-energy regime. A molecular mechanism accounting for the desorption event remains elusive.

Photon-stimulated desorption is a powerful tool to study fundamental processes in adsorbate-surface systems, as well as to achieve selective surface reactions for controlled surface processing (*1*–*4*). Photons are easily directed and tuned in energy to induce transitions in atomic and molecular states, with high spatial and temporal precision. Direct adsorbate-surface bond breaking by electronic excitation using ultraviolet light has been reported (*5*, *6*). However, visible and infrared (IR)–stimulated desorption processes studied so far generally involve indirect mechanisms (*7*, *8*), such as light-induced substrate heating (*9*) and, in physisorbed systems, energy transfer from internal molecular excitation to molecular translational motion away from the surface (*10*, *11*). Selective bond scission at these lower energies is desirable but has proven challenging because of rapid energy delocalization from the mode of excitation (*1*). Here we report resonant photodesorption of H from a Si(111) surface using IR radiation. We show that the process is resonant with the Si-H vibrational energy and displays an unusual and surprising dependence on excitation intensity, which cannot be explained by either thermal or electronic processes. Successful elucidation of this fundamental excitation mechanism would be a major advance in surface science, and its implementation could lead to site-selective desorption at low temperatures.

The H/Si(111) structure is a well-characterized adsorbate system, ideal for the study of fundamental surface processes. The Si-H bond is perpendicular to the Si(111) surface, with a bond energy estimated to be between 3.15 and 3.35 eV (*12*) and a vibrational stretch energy of 0.26 eV at the terrace sites (*13*). In our experiments, the Si-H vibrational stretch mode was resonantly excited by IR photons. The Si substrate was transparent to mid-IR illumination, minimizing electronic excitation and direct laser-induced heating. Although this paper deals with the basic mechanisms of the desorption process, there could be important technological applications. Because H desorption is an important component of Si chemical vapor deposition, modifying the H desorption mechanism can have a large impact on growth. In particular, the vibrational energy of an Si-H bond at the terrace site of the Si(111) surface differs from that at the step site by 51.8 cm^{-1} (*14*), opening the possibility of site-selective desorption of H adatoms by IR irradiation. Such a photolytic process could efficiently treat a large area and modify the type of sites available for epitaxial growth. This might be compared to nanoscale lithography of H on Si achieved using the scanning tunneling microscope (STM) (*15*, *16*).

Our experiment was performed in an ultrahigh-vacuum (UHV) system at room temperature. The base pressure was $\sim 3.0 \times 10^{-10}$ Torr. The sample was prepared in air, using hydrofluoric acid etching to protect the Si surface from oxidation and contamination, and then cleaned by direct-current heating in the UHV chamber. The clean surface was verified by the strong Si(111)-7×7 low-energy electron diffraction (LEED) pattern. Ultrahigh-purity H_2 gas was introduced into the UHV chamber. A tungsten filament at 2000 K was used to dissociate H_2 molecules, thus efficiently forming

[1]Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455, USA. [2]Department of Physics and Astronomy, Vanderbilt University, Nashville, TN 37235, USA. [3]Materials Science and Technology Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA. [4]Department of Physics and Astronomy, University of Tennessee, Knoxville, TN 37996, USA.

*To whom correspondence should be addressed. E-mail: picohen@umn.edu

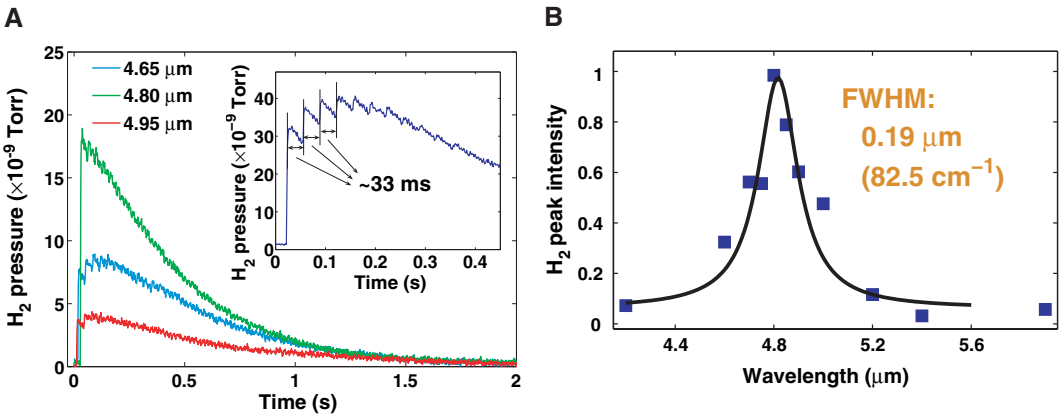


Fig. 1. (**A**) Typical curves of desorbed H_2 signal as a function of time at three wavelengths. The background pressure has been subtracted. The FEL fluence was 0.8 J/cm^2 per macropulse. The inset shows the fine structure of a desorption curve. The step rises are due to desorption by individual FEL macropulses. The 33-ms interval corresponds to the 30-Hz repetition rate of the macropulse. (**B**) Wavelength dependence of the desorption yield. Data were normalized to 1. The solid curve is a Lorentzian fit to the data.

the Si-H bonds at the Si(111) surface. After a hydrogen dose of 3×10^{-6} Torr for 120 s, an area of Si surface measuring 5 by 4 mm was uniformly covered with H atoms, as evidenced by the Si(111)-1×1 LEED pattern and the uniform desorption behavior at different spots when all other desorption conditions were kept the same. Thermal desorption spectra showed only one peak at ~820 K, indicating that H was adsorbed exclusively as a monohydride phase on the Si surface (*17*).

The free-electron laser (FEL) at Vanderbilt University provides a maximum of 100 mJ per macropulse at a repetition rate of 30 Hz. Each macropulse is composed of ~10^4 micropulses of 1 ps duration, spaced 350 ps apart. The spectral width of the FEL emission is ~50 to 90 cm^{-1}. The IR radiation was guided into the chamber through a CaF_2 viewport and focused onto the sample with a spot size of 0.8 mm. Its power and polarization could be adjusted with an attenuator and polarizer outside the chamber. The sample surface was positioned to make a 30° angle with the linearly polarized beam. The total electric field above the Si surface is the sum of the incident and the reflected fields. For the Si substrate with refractive index 3.42 and this incident angle of 60°, the direction of this total electric field is close to the surface normal; namely, the direction of the Si-H bond. The field projection in this direction is $(1 + r_p)E\cos\theta\cos 30°$, with E being the incident field, r_p the reflection coefficient of the p-polarized component, and θ the angle between the FEL electric field and the incident plane. The effective FEL fluence in the direction of the Si-H bond can be varied by changing either the incident intensity E^2 or the polarization angle θ.

The sample was exposed to the FEL illumination at room temperature, and the desorbed H was pumped away by the ion pump at the same time. An IR pyrometer focused on the sample detected no temperature rise during desorption, showing no FEL-induced heating of the bulk sample, though not ruling out a local transient temperature increase. A quadrupole mass analyzer was employed to record the H_2 partial pressure as a function of time once the FEL shutter was open. After desorption, an experimental parameter, such as the FEL wavelength or polarization, was varied, and the FEL focus was moved to a new sample position that had not been exposed to the FEL illumination. When the FEL beam was not incident on the Si surface, no desorption signal was measured, indicating that there was no contribution from scattered FEL radiation.

To study the wavelength dependence of H_2 photodesorption, we monitored the H_2 partial pressure as a function of time over a range of FEL wavelengths, with the fluence kept fixed and the polarization set to θ = 0° (Fig. 1A). The signal quickly rose to its peak, then decayed exponentially with a rate that was in agreement with the pumping speed of the vacuum system. The fine structure of the desorption curves (inset, Fig. 1A) features step-like increases. The interval of 33 ms is in good agreement with the 30-Hz repetition rate of the FEL macropulse, indicating that these increases resulted from desorption by each FEL macropulse. Finer temporal structure due to FEL micropulses could not be resolved by our instrumentation. The H_2 partial pressure is a measure of the total number of H_2 molecules in the UHV chamber. Therefore, the magnitude of each jump is taken as a measure of the number of H_2 molecules desorbed by each macropulse; that is, the desorption yield.

The desorption yield peaked at a wavelength of 4.8 μm (Fig. 1B), corresponding to 0.26 eV, the energy of the vibrational stretch mode of the Si-H bond at the terrace sites of the Si(111) surface. A Lorentzian fit yielded a full width at half maximum (FWHM) of 0.19 μm (82.5 cm^{-1}), close to the FEL linewidth, which is the limiting factor in resolving wavelength. Because the vibrational energy of the Si-H bond at the terrace site differs from that at the step site by 51.8 cm^{-1}, this resonance effect suggests that wavelength-selective site desorption (*18*) would be readily possible with an incident beam of appropriately narrow linewidth. In addition, desorption was unobservable when the wavelength was far from resonance, ruling out a desorption mechanism due to simple, direct laser heating.

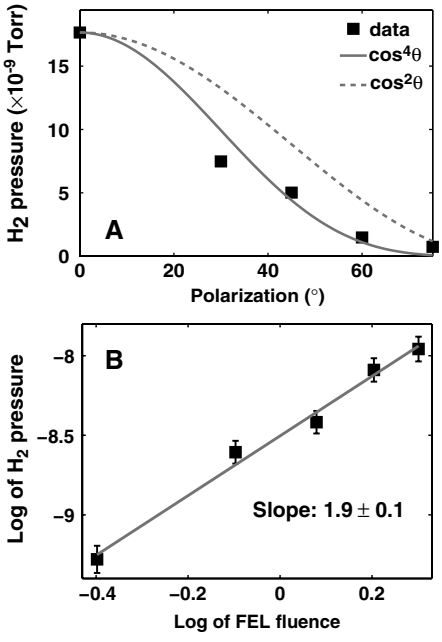


Fig. 2. (**A**) Polarization dependence of the desorption yield (squares). The FEL fluence was kept at 0.8 J/cm^2 per macropulse for all polarizations. (**B**) Log-log plot of desorbed H_2 versus FEL fluence. A linear fit yields a slope of 1.9 ± 0.1. The polarization was 0° for all fluences, and the wavelength was tuned to 4.8 μm.

A study of the polarization dependence of the desorption yield provides further insight into the desorption mechanism. For bulk-like thermal processes, possibly because of impurities or defects, the desorption yield is not expected to be very sensitive to the polarization angle. For direct bond-breaking processes via electronic excitation, the yield should be a simple quadratic function of the FEL electric field projected in the direction of Si-H bond (*5*) (that is, a linear dependence on the incident intensity) and follow a $\cos^2\theta$ curve as illustrated (Fig. 2A). In our experiment with constant FEL fluence, however, the desorption yield scales with the incident polarization as $\cos^4\theta$ (Fig.

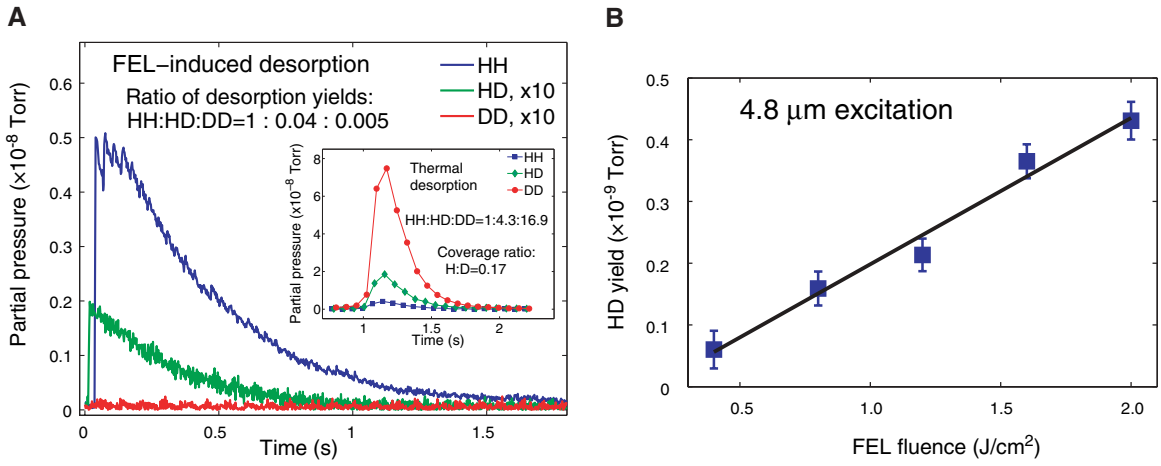


Fig. 3. (**A**) Partial pressures of H_2, HD, and D_2 as a function of time arising from FEL-induced desorption. The fluence was 1.6 J/cm^2 per macropulse. The FEL data for HD and D_2 are magnified by 10 times. The inset shows the partial pressures in a single thermal desorption. Data were obtained by continuous cycling of an atomic mass unit scan from 0.5 to 5.5, with a period of 75 ms. The temperature began to rise at 0.9 s, from room temperature to the terminal temperature of ~950°C. (**B**) Desorption yield of HD as a function of FEL fluence when excited at 4.8 μm. A linear relationship is observed.

2A). This behavior indicates that the observed desorption does not result from a thermal process caused by direct laser heating, but from the direct interaction between the incident field and the Si-H bonds. Furthermore, the $\cos^4\theta$ dependence implies a quadratic dependence on the incident intensity, as verified in a measurement of the yield dependence on the FEL fluence (Fig. 2B). This quadratic dependence is significantly different from the ~ninth-order dependence observed in multiple vibrational excitations of H on Si(100) by the STM (*19*).

To further probe the nature of the desorption mechanism, the thermal desorption and IR-induced desorption of a coadsorbed mixture of H and D were also measured. H and D adatoms were coadsorbed and then the desorption yields of different species were measured (Fig. 3A). For thermal desorption, the ratio of the yields was H_2:HD:D_2 = 1:4.3:16.9. In the FEL-induced process, with the same H/D coverage and the wavelength tuned to the Si-H stretch at 4.8 μm, the desorption ratio was found to be 1:0.04:0.005, in stark contrast to the thermal data. Furthermore, the HD desorption yield was linear in photon flux (Fig. 3B) over the same range in which the H_2 yield was seen to be quadratic. These results support the idea that there are very different mechanisms for thermal desorption and IR-induced desorption.

The underlying mechanism that may account for the observed desorption behavior is not immediately obvious. Related phenomena include laser-induced thermal desorption and vibrational ladder-climbing. Direct laser heating of the substrate (*9*) can clearly be ruled out as the dominant mechanism, based on the wavelength and polarization dependencies of the desorption yield and our experimental measurement of the bulk temperature. Another type of thermal mechanism, resonant local heating, refers to a process in which the Si-H vibrational stretch mode is resonantly excited by the IR photons and then couples to local substrate phonons, causing a local temperature increase and eventually thermal desorption. The measured isotope effect in this FEL-induced desorption rules out such local heating, because the FEL-induced desorption yields are drastically different from thermal desorption yields. In the vibrational ladder-climbing mechanism, Si-H bonds may be continuously excited to higher vibrational states by either multiphoton adsorption (*10*) or dipole-dipole coupling of Si-H bonds (*20*). We estimate that an H adatom must climb to approximately the fifth vibrational energy level of the Si-H stretch mode in order to overcome either the diffusion barrier of 1.5 eV (*21*) for enhanced diffusion or half of the activation energy of 2.4 eV (*22*) in a thermal activation scheme. The quadratic power dependence observed in the H_2 desorption and the linear dependence in the HD desorption might not necessarily suggest a two-photon or one-photon process. Multiphoton processes often display a much lower order of dependence on the photon flux than the number of photons absorbed (*23*). For example, if the absorption cross section increases when the system has absorbed one or two photons, the excitation rate will be determined by these initial photons. Consequently, the measured power dependence will be low-order, even though the process requires additional photons. In such a multiphoton process, the higher energy states would have slightly different absorption energies because of anharmonic effects. However, the FEL linewidth is broad enough to allow the excitation of some number of levels. As a result, these considerations make vibrational ladder-climbing a possible mechanism to explain the experimental observations. New mechanisms that are consistent with the data might also be considered and further investigated.

The experimental findings reported above are unprecedented and unexpected. The precise underlying mechanism is still to be identified. An important intention of this work is to stimulate future studies on the basic mechanisms of surface dynamics, both experimentally and theoretically.

References and Notes

1. T. J. Chuang, *Surf. Sci. Rep.* **3**, 1 (1983).
2. X.-Y. Zhu, *Annu. Rev. Phys. Chem.* **45**, 113 (1994).
3. P. A. Williams *et al.*, *Phys. Rev. Lett.* **79**, 3459 (1997).
4. M. Dürr, A. Biedermann, Z. Hu, U. Höfer, T. F. Heinz, *Science* **296**, 1838 (2002).
5. T. Vondrak, X.-Y. Zhu, *Phys. Rev. Lett.* **82**, 1967 (1999).
6. L. T. Hudson *et al.*, *Phys. Rev. B* **62**, 10535 (2000).
7. Ph. Avouris, R. E. Walkup, *Annu. Rev. Phys. Chem.* **40**, 173 (1989).
8. F. M. Zimmermann, W. Ho, *Surf. Sci. Rep.* **22**, 127 (1995).
9. B. G. Koehler, S. M. George, *Surf. Sci.* **248**, 158 (1991).
10. J. Heidberg, H. Stein, E. Riehl, *Phys. Rev. Lett.* **49**, 666 (1982).
11. Z. W. Gortel, H. J. Kreuzer, P. Piercy, R. Teshima, *Phys. Rev. B* **28**, 2119 (1983).
12. C. Herring, C. G. Van de Walle, *Phys. Rev. B* **55**, 13314 (1997).
13. K. Raghavachari, P. Jakob, Y. J. Chabal, *Chem. Phys. Lett.* **206**, 156 (1993).
14. P. Jakob, Y. J. Chabal, *J. Chem. Phys.* **95**, 2897 (1991).
15. E. J. Buehler, J. J. Boland, *Science* **290**, 506 (2000).
16. J. W. Lyding *et al.*, *Appl. Surf. Sci.* **130**, 221 (1998).
17. P. Gupta, V. L. Colvin, S. M. George, *Phys. Rev. B* **37**, 8234 (1988).
18. B. Wu, P. I. Cohen, L. C. Feldman, Z. Zhang, *Appl. Phys. Lett.* **84**, 2175 (2004).
19. T.-C. Shen *et al.*, *Science* **268**, 1590 (1995).
20. S. A. Corcelli, J. C. Tully, *J. Phys. Chem. A* **106**, 10849 (2002).
21. G. A. Reider, U. Höfer, T. F. Heinz, *Phys. Rev. Lett.* **66**, 1994 (1991).
22. G. A. Reider, U. Höfer, T. F. Heinz, *J. Chem. Phys.* **94**, 4080 (1991).
23. T. J. Chuang, *J. Chem. Phys.* **72**, 1453 (1981).

24. We thank B. N. J. Persson, J. C. Tully, M. M. Albert, B. Wu, J. Rozen, and S. Meng for helpful discussions and J. Kozub, B. Gabella, and the personnel of the Vanderbilt University FEL Center for providing technical support. This work is funded by Defense Advanced Research Projects Agency grant N66001-04-1-8924; U.S. Department of Energy (DOE) grant ER45781; NSF grant DMR-0306239; and in part by the Laboratory-Directed Research and Development Program of Oak Ridge National Laboratory, managed by UT-Battelle for DOE under grant DE-AC05-00OR22725.

3 January 2006; accepted 17 March 2006
10.1126/science.1124529

Oligonucleotide-Modified Gold Nanoparticles for Intracellular Gene Regulation

Nathaniel L. Rosi,* David A. Giljohann,* C. Shad Thaxton, Abigail K. R. Lytton-Jean, Min Su Han, Chad A. Mirkin†

We describe the use of gold nanoparticle-oligonucleotide complexes as intracellular gene regulation agents for the control of protein expression in cells. These oligonucleotide-modified nanoparticles have affinity constants for complementary nucleic acids that are higher than their unmodified oligonucleotide counterparts, are less susceptible to degradation by nuclease activity, exhibit greater than 99% cellular uptake, can introduce oligonucleotides at a higher effective concentration than conventional transfection agents, and are nontoxic to the cells under the conditions studied. By chemically tailoring the density of DNA bound to the surface of gold nanoparticles, we demonstrated a tunable gene knockdown.

Nucleic acid-based methods for controlling gene expression have had a considerable impact on research involving gene pathways and function (*1*–*3*). In addition, antisense (AS) therapies are potentially powerful candidates for clinical treatments of various ailments, including cancer and HIV/AIDS (*1*, *4*). In conventional AS approaches, oligonucleotides designed to hybridize with target mRNA sequences downregulate the expression of the corresponding translated proteins. Although the potential of AS oligodeoxyonucleotides (ASODNs) was recognized more than 20 years ago (*5*, *6*), their development into viable therapeutic systems has faced challenges with regard to stable transfection and entry into diverse cell types, toxicity, and low efficacy (*3*). Various transfection agents have been developed to shuttle nucleic acids into cells, including cationic lipids and polymers, modified viruses, dendrimers, liposomes, and nanoparticles (*1*, *4*, *7*, *8*), but each system has particular limitations.

Gold nanoparticles (Au NPs) have a high affinity for biomolecules and have been used as nonviral vectors for DNA delivery (*9*–*13*). Gold nanoparticles chemically functionalized with alkylthiol-terminated oligonucleotides (*14*) are highly stable in saline solutions and bind complementary nucleic acids in a very selective and cooperative manner (*15*); their equilibrium association constants can be more than two orders of magnitude greater than those observed for unmodified oligonucleotides and their complements (*16*). These distinctive properties have made oligonucleotide-functionalized Au NPs the centerpiece of several highly sensitive and selective assays for biomolecule detection (*17*–*21*). Because of their demonstrated stability and enhanced binding properties, we hypothesized that this particular type of Au NP could efficiently scavenge intracellular DNA or RNA. We introduce the use of oligonucleotide-functionalized Au NPs as new antisense agents that take advantage of the ensemble properties of the nanoparticle-oligonucleotide conjugate.

In initial experiments, we targeted mRNA sequences coding for enhanced green fluorescent protein (EGFP) expressed in C166, a mouse endothelial cell line (ATCC). A phosphorothioate-modified antisense sequence complementary to an internal coding region (bases 1198 to 1215) of the mRNA for EGFP was selected from published literature (*22*). This sequence was used in the design of two sets of antisense Au NPs, with the ASODN conjugated to the Au NP surface with either one or four thiol groups (Fig. 1). The tetra-thiol particle (particle A) supports 45 to 50 strands, whereas the mono-thiol particle (particle B) has 110 to 120 strands. Whereas particle A exhibits a binding constant to its complementary sequence that is approximately equal to that of an unmodified ASODN, particle B has an affinity ~35 times as high as that of the unmodified ASODN (fig. S1) (*23*). This result is consistent with a cooperative binding theory, which predicts that higher oligonucleotide packing densities result in a corresponding increase in association constant (*15*). Taken together, particles A and B offer the opportunity to study the potential of ASNPs in regulating gene expression and, more specifically, the effect of particle binding constants and oligonucleotide loading on the performance of such particles in the context of EGFP expression.

We evaluated the Au NP-ASODN composites (ASNPs) for their cellular uptake, intracellular stability, and the fate of the oligonucleotides once the particles were taken up by the cell. Most approaches to transfection agents rely on the use of positively charged materials to ensure cellular uptake (*9*, *24*). Unexpectedly, ASNPs readily enter cells despite their coating with negatively charged DNA. Entry of the ASNPs into the EGFP-expressing C166 mouse endothelial cells was confirmed by incubating the cells for 48 hours in the presence of ASNPs functionalized with 5′ Cy5.5-labeled

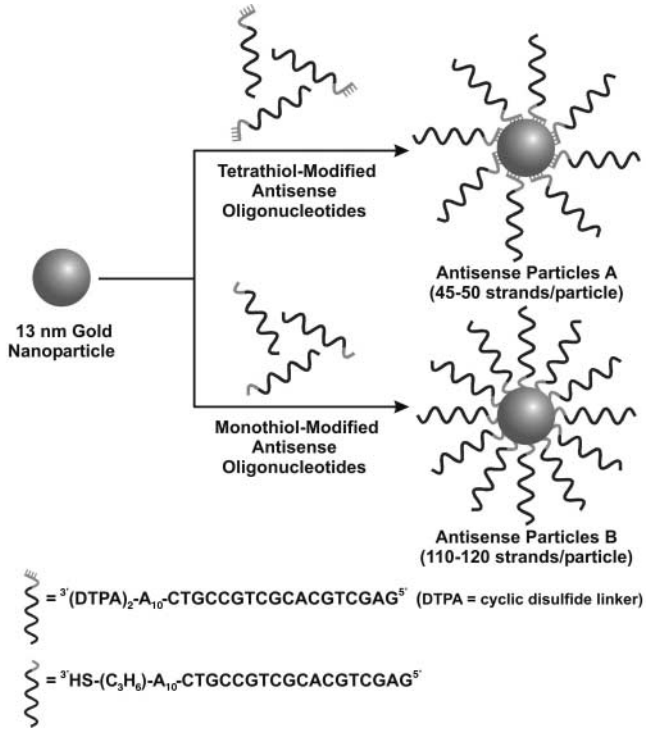


Fig 1. Preparation of antisense Au NPs. Citrate-stabilized gold nanoparticles (13 nm ± 1 nm) were functionalized with ASODNs that were premodified with an A_{10} tether and either two cyclic disulfides (DTPA) or an alkyl-thiol anchoring group to produce Antisense Particles A and B, respectively.

Department of Chemistry and International Institute for Nanotechnology, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208–3113 USA

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: chadnano@northwestern.edu

ASODNs. The uptake was studied using confocal fluorescence microscopy. Observation of Cy5.5 fluorescence throughout the cytoplasm provided proof of particle uptake. Additionally, we performed identical uptake experiments with various cell types, including RAW 264.7 (macrophage), HeLa (cervical carcinoma), NIH-3T3 (fibroblast), and MDCK (kidney) (fig. S2). In each case, we observed >99% of uptake (virtually every cell incorporated the ASNPs), and there were no differences in cellular morphology as compared with untreated control cells or cell viability under the conditions studied as determined by Trypan Blue staining.

We determined whether the ASODNs remain bound to the gold particle surface within the cell by using "dual fluorophore" ASODNs labeled with both a 3′ fluorophore (Cy3, on the surface of the particle) and a 5′ fluorophore (Cy5.5, at the end of the ASODN) [see (*23*) for sequence and preparation]. If ASODNs remain stably attached to the particle surface within the cell, both their Cy3 and Cy5.5 fluorescence would be quenched. If they were digested by nucleases, we expected to see free Cy5.5, whereas if they were displaced from the surface of the particle by chemical reduction, we expected to see strong emission from both dyes.

After a 48-hour incubation period in the presence of ASNPs, two-photon confocal laser scanning microscopy of cells (Fig. 2A) with both 546 nm and 633 nm light resulted only in fluorescence from the 5′ Cy5.5 fluorophore, which could be attributed to some nuclease digestion of the ASODNs and residual fluorescence due to incomplete quenching by the Au NPs (*25*) (the Cy5.5 is approximately 9 nm from the gold surface). Because only negligible 3′ Cy3 fluorescence was observed, we conclude that the ASODNs do, in fact, remain chemically attached to the Au NP surface while inside the cell. In additional experiments, C166 cells were incubated with ASNPs functionalized with 5′ Cy5.5-modified ASODNs. After a 48-hour incubation period, the cells were lysed and the bulk fluorescence of the lysate was measured in the emission range of the Cy5.5 fluorophore (706 to 717 nm). Dithiothreitol (DTT) displacement of the Cy5.5-labeled ASODNs from the nanoparticles resulted in increased Cy5.5 fluorescence emission in the lysate, indicating that many of the ASODNs remain undigested by nucleases after 48 hours within the cellular environment (fig. S4).

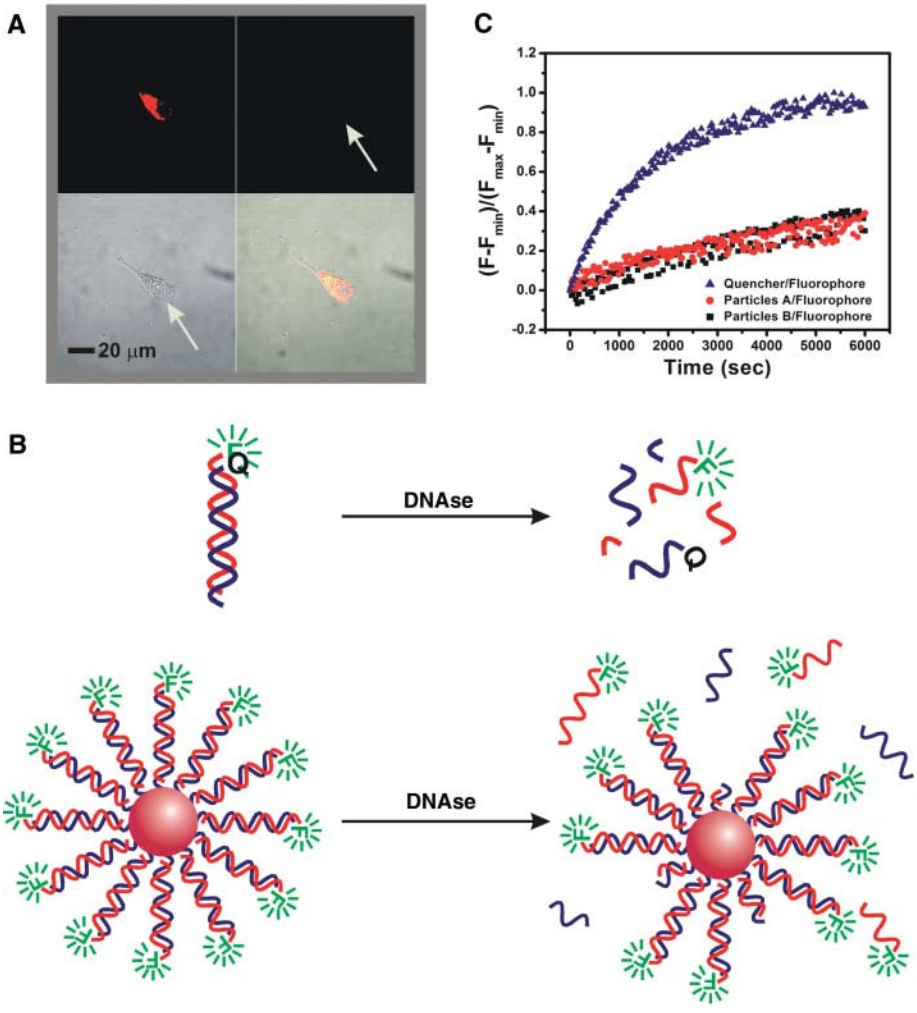


Fig. 2. Experiments aimed at understanding the intracellular stability of antisense nanoparticles. (**A**) Fluorescent microscopy images of C166-EGFP cells incubated 48 hours with antisense particle B functionalized with dual-fluorophore labeled ASODNs (3′ Cy3 and 5′ Cy5.5) only reveal fluorescence from Cy5.5 (706 to 717 nm, upper left). Negligible fluorescence is observed in the emission range of Cy3 (565 to 615 nm, upper right). Transmission and composite overlay images are shown in the lower left and lower right quadrants, respectively. The arrows indicate the location of the cell. Similar data collected from experiments using particle A are included in fig. S3 (*23*). (**B**) Duplexes composed of either quencher-modified ASODN/fluorophore-modified complement or antisense particle/fluorophore-modified complement were treated with DNase. (**C**) The ASODN duplex degraded much faster than the antisense particle duplex, as calculated using fluorescence spectroscopy, where F_{min} is the fluorescence of the mixture at its initial, fully hybridized state and F_{max} is the maximum fluorescence of the system at 80°C where all of the oligonucleotides are dehybridized.

Glutathione (GSH) has been used as a reagent to effect DNA release from NPs in the context of gene delivery (*26*). Intracellular GSH levels differ between various cell types (*27*), so GSH levels in C166 cells could be lower than other cell lines. However, analysis of the stability of the dual fluorophore-labeled DNA on the Au NP (*23*) shows that even at elevated GSH levels (10 mM), >60% of the oligonucleotides remain bound to the particle after 48 hours. For cellular systems in which GSH levels are high, it is possible to increase the strength of the interaction of ASODN with Au NPs through the use of cyclic disulfide anchoring groups without substantially changing surface coverage (*28*, *29*). In the case of C166 cells studied, we can conclude from the dual-fluorophore particle stability experiments that the oligonucleotides remain attached to the particle surface.

Given the potential for ASODN degradation by nucleases in vivo, we next examined the in vitro resistance of the ASNP to DNase compared with particle-free oligonucleotides. Fluorescence experiments indicate that particle-bound ASODNs are degraded much slower than are unbound ASODNs (Fig. 2, B and C). The tight packing of the ASODNs on the NP surface likely causes steric inhibition of nuclease degradation. Others have made similar observations in the context of assembly and manipulation of oligonucleotide-Au NP superstructures (*30*). The increased resistance to nuclease degradation of particle-bound ASODNs should increase the lifetime of the AS agent within the cell.

The activity of the ASNPs with respect to gene knockdown was evaluated with EGFP-expressing C166 cells. Evidence of EGFP knockdown was first observed by two-photon excitation confocal fluorescence microscopy, which indicated that cells treated with ASNPs exhibited noticeably lower fluorescence compared with untreated control cells (Fig. 3). These

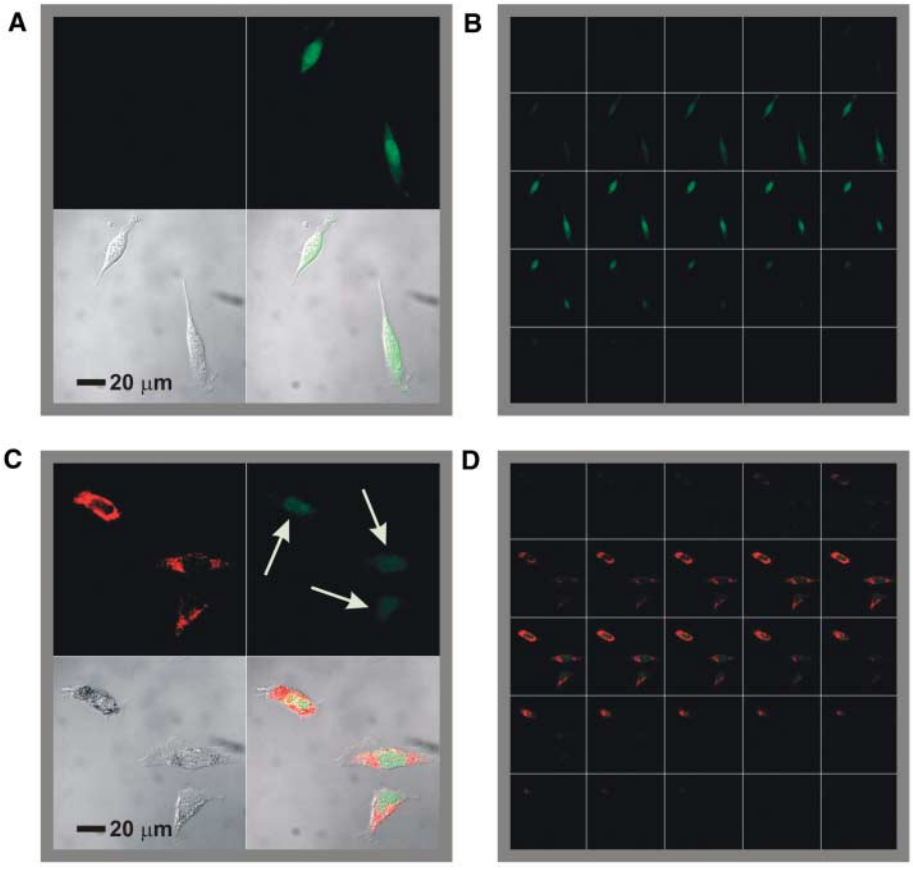


Fig. 3. Confocal fluorescence microscopy images showing EGFP knockdown. (A) Untreated control C166-EGFP cells (upper left, Cy5.5 emission, 706 to 717 nm; upper right, EGFP emission, 500 to 550 nm; lower left, transmission image of cells; lower right, composite overlay of all three channels) showed a significant amount of emission throughout the cell. (B) 1 μm sectioning images of control cells. (C and D) Cells treated with antisense particles A or B showed a decrease in the amount of EGFP emission.

observations were then confirmed by quantitatively assaying for EGFP expression with fluorescence spectroscopy.

As previously described, the binding properties of the ASNPs can be tailored by controlling the number of ASODNs loaded on the particles. We expected that particle B, with its higher binding affinity, would cause greater reductions in EGFP expression than particle A. To test this hypothesis, samples of particle A were added to C166-EGFP cells. After 48 hours, the cells were collected, lysed, and assayed for EGFP expression [see (*23*) for quantification methods)]. Noncomplementary ODN-Au NPs (*23*) were used as controls and showed similar fluorescence levels to nontreated cells (Table 1). Cells incubated with 0.024 nmol or 0.048 nmol of particle A (low binding constant) displayed a 12% and 14% decrease in EGFP expression, respectively. However, when incubated with the same amount of particle B (tight binding constant), the cells displayed a decrease in EGFP expression of 14% and 20%, respectively (Table 1). The differences in the observed knockdown between Particles A and B indicate that the ASNP binding constant can significantly influence the amount of protein expression.

Interestingly, when the particles were compared with commercially available lipoplexes (Lipofectamine 2000, Invitrogen; Cytofectin, Gene Therapy Systems) in the context of EGFP expression, the nanoparticle system outperformed the commercial systems under the conditions studied with respect to percent knockdown, total amount of ASODN delivery, and nontoxicity (Table 1). Following recommended commercial protocols, Lipofectamine and Cytofectin were used to transfect EGFP-C166 cells with an equal

Table 1. Performance characteristics of antisense nanoparticles.

Antisense experiment	Observed toxicity	Approximate binding constant	Percent decrease in EGFP expression
Antisense Particles A (0.024 nmol particles 1.08 nmol ASODN)	No	7.1×10^{20}	11 ± 2
Antisense Particles A (0.048 nmol particles 2.16 nmol ASODN)	No	7.1×10^{20}	14 ± 0.4
Antisense Particles B (0.024 nmol particles 2.64 nmol ASODN)	No	2.6×10^{22}	14 ± 1
Antisense Particles B (0.048 nmol particles 5.28 nmol ASODN)	No	2.6×10^{22}	20 ± 4
Nonsense Particles A (0.048 nmol particles)	No	N/A	0 ± 3
Nonsense Particles B (0.048 nmol particles)	No	N/A	0 ± 2
Lipofectamine 2000 (0.024 nmol ASODN)	No	6.7×10^{20}	6 ± 0.2
Lipofectamine 2000 (2.64 nmol ASODN)	Yes	6.7×10^{20}	N/A
Cytofectin (0.024 nmol ASODN)	No	6.7×10^{20}	7 ± 0.7
Cytofectin (2.64 nmol ASODN)	Yes	6.7×10^{20}	N/A

amount of ASODNs delivered using antisense nanoparticles (0.024 nmol ASNP: 2.64 nmol ASODN); however, extreme toxicity was observed in both cases as measured by cell death, so we lowered the amount of ASODN and transfection reagent to the point that one strand of transfected ASODN was equivalent to one antisense particle (0.024 nmol ASNP: 0.024 nmol ASODN). In these cases, ASODNs transfected with either Lipofectamine or Cytofectin resulted in only ~6 to 8% knockdown in EGFP expression (Table 1). Evaluation of multiple cell lines (*23*) to determine toxicity limits of the ODN-Au NPs have not shown any appreciable cell death even with high (0.12 nmol) particle loadings.

Although these systems have not yet been optimized for maximum efficacy, the ability to systematically control the oligonucleotide loading on the NP surface has allowed us to identify several features that make them attractive candidates for antisense studies and therapies. The ability to modify the gold nanoparticle surface allows one to realize unusual cooperative properties that lead to enhanced target binding and allows the introduction of a variety of functional groups that have proven to be informative in terms of studying how the structure works within a cell. Moreover, this platform will allow us and others to add functionality (*31*, *32*) that could direct the oligonucleotide-modified nanoparticle agents to specific cell types and different components within the cell, thus opening the door for new possibilities in the study of gene function and nanotherapies.

References and Notes

1. S. D. Patil, D. G. Rhodes, D. J. Burgess, *AAPS J.* **7**, E61 (2005).
2. M. T. McManus, P. A. Sharp, *Nat. Rev. Genet.* **3**, 737 (2002).
3. I. Lebedeva, C. A. Stein, *Annu. Rev. Pharmacol. Toxicol.* **41**, 403 (2001).
4. T. L. H. Jason, J. Koropatrick, R. W. Berg, *Toxicol. Appl. Pharm.* **201**, 66 (2004).
5. M. L. Stephenson, P. C. Zamecnik, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 285 (1978).
6. P. C. Zamecnik, M. L. Stephenson, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 280 (1978).
7. D. J. Bharali *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 11539 (2005).
8. A. U. Bielinska, C. L. Chen, J. Johnson, J. R. Baker, *Bioconjugate Chem.* **10**, 843 (1999).
9. M. Thomas, A. M. Klibanov, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9138 (2003).
10. P. Sundaram, W. Xiao, J. L. Brandsma, *Nucleic Acids Res.* **24**, 1375 (1996).
11. K. K. Sandhu, C. M. McIntosh, J. M. Simard, S. W. Smith, V. M. Rotello, *Bioconjugate Chem.* **13**, 3 (2002).
12. M. M. Ow Sullivan, J. J. Green, T. M. Przybycien, *Gene Ther.* **10**, 1882 (2003).
13. C-.P. Jen *et al.*, *Langmuir* **20**, 1369 (2004).
14. C. A. Mirkin, R. L. Letsinger, R. C. Mucic, J. J. Storhoff, *Nature* **382**, 607 (1996).
15. R. Jin, G. Wu, Z. Li, C. A. Mirkin, G. C. Schatz, *J. Am. Chem. Soc.* **125**, 1643 (2003).
16. A. K. R. Lytton-Jean, C. A. Mirkin, *J. Am. Chem. Soc.* **127**, 12754 (2005).
17. N. L. Rosi, C. A. Mirkin, *Chem. Rev.* **105**, 1547 (2005).
18. R. Elghanian, J. J. Storhoff, R. C. Mucic, R. L. Letsinger, C. A. Mirkin, *Science* **277**, 1078 (1997).
19. J.-M. Nam, C. S. Thaxton, C. A. Mirkin, *Science* **301**, 1884 (2003).
20. L. He *et al.*, *J. Am. Chem. Soc.* **122**, 9071 (2000).
21. D. J. Maxwell, J. R. Taylor, S. M. Nie, *J. Am. Chem. Soc.* **124**, 9606 (2002).
22. J-R. Bertrand, M. Pottier, A. Vekris, P. Opolon, A. Maksimenko, C. Malvy, *Biochem. Biophys. Res. Commun.* **296**, 1000 (2002).
23. Materials and methods are available as supporting material on *Science* Online.
24. A. G. Tkachenko *et al.*, *Bioconjugate Chem.* **15**, 482 (2004).
25. E. Dulkeith, A. C. Morteani, T. Niedereichholz, T. A. Klar, J. Feldmann, *Phys. Review Letters* **89**, 20 203002-1 (2002).
26. Han *et al.*, *Bioconjugate Chem.* **16**, 1356 (2005).
27. Anderson *et al.*, *Chem.-Biol. Interact.* **112**, 1 (1998).
28. R. L. Letsinger, R. Elghanian, G. Viswanadham, C. A. Mirkin, *Bioconjugate Chem.* **11**, 289 (2000).
29. Z. Li, R. Jin, C. A. Mirkin, R. L. Letsinger, *Nucleic Acids Res.* **30**, 1558 (2002).
30. Z. Wang, A. G. Kanaras, A. D. Bates, R. Cosstick, M. Brust, *J. Mater. Chem.* **14**, 578 (2004).
31. J. M. Bergen, S. H. Pun, *MRS Bull.* **30**, 663 (2005).
32. R. Weissleder, K. Kelly, E. Y. Sun, T. Shtatland, L. Josephson, *Nat. Biotechnol.* **23**, 1418 (2005).

33. C.A.M. gratefully acknowledges a NIH Director's Pioneer Award and a National Cancer Institute–Center of Cancer Nanotechnology Excellence grant for support of this research. The authors thank T. O'Halloran, T. Meade, L. Shea, V. Band, H. Godwin, C. B. Gurumurthy, R. I. Morimoto, W. Russin, The Biological Imaging Facility, and the Keck Biophysics Facility for assistance.

Supporting Online Material

www.sciencemag.org/cgi/content/full/312/5776/1027/DC1
Materials and Methods
Figs. S1 to S4

30 January 2006; accepted 5 April 2006
10.1126/science.1125559

Preparation of Poly(diiododiacetylene), an Ordered Conjugated Polymer of Carbon and Iodine

Aiwu Sun, Joseph W. Lauher, Nancy S. Goroff*

Conjugated organic polymers generally must include large substituents for stability, either contained within or appended to the polymer chain. In polydiacetylenes, the substituents fulfill another important role: During topochemical polymerization, they control the spacing between the diyne monomers to produce an ordered polymer. By using a co-crystal scaffolding, we have prepared poly(diiododiacetylene), or PIDA, a nearly unadorned carbon chain substituted with only single-atom iodine side groups. The monomer, diiodobutadiyne, forms co-crystals with bis(nitrile) oxalamides, aligned by hydrogen bonds between oxalamide groups and weak Lewis acid-base interactions between nitriles and iodoalkynes. In co-crystals with one oxalamide host, the diyne undergoes spontaneous topochemical polymerization to form PIDA. The structure of the dark blue crystals, which look copper-colored under reflected light, has been confirmed by single-crystal x-ray diffraction, ultraviolet-visible absorption spectroscopy, and scanning electron microscopy.

Since Heeger, McDiaramid, and Shirakawa's discovery of conducting behavior in doped polyacetylene samples (*1*), carbon-rich molecules and materials have attracted great interest for their electronic and optical properties. Recently examined applications include photovoltaic cells, organic light-emitting diodes (OLEDs), field-effect transistors, and chemical sensors (*2*–*5*). Polyacetylene, the simplest conjugated polymer, consists of a backbone of carbon atoms, each bonded to one hydrogen atom and connected together by alternating single and double bonds. However, this material is thermally unstable and insoluble, making it unsuitable for general use. More manageable conjugated polymers can be prepared from substituted alkynes (*6*) and other conjugated monomers, such as substituted aromatics; for example, in the commercially useful polymer poly(*p*-phenylenevinylene) (PPV), the backbone contains benzene rings as well as simple carbon-carbon double bonds. Polydiacetylenes, in which every other double bond in the polyacetylene backbone is replaced by a triple bond, are potent multiphoton absorbers that have applications in optical limiters, waveguides, and thermometric sensors (*2*).

Obtaining the desired optical and electronic properties of a given conjugated polymer requires a highly ordered molecular structure. For polydiacetylenes, this presents a substantial synthetic challenge: With solution-phase methods, the monomer diynes can react in either a 1,2- or 1,4-fashion, leading to a polymer with an irregular, branched backbone instead of regular repeat units. However, as discovered by Wegner (*7*) and elaborated by Baughman (*8*), if the diynes are first aligned appropriately in the solid state at a distance commensurate with the repeat distance in the target polymer, then their arrangement in space will control their reactivity, leading to ordered topochemical polymerization. Chemists have developed many creative approaches to induce appropriate alignment for 1,4-polymerization, not only in the solid state, but also in Langmuir-Blodgett films, liposomes and vesicles, and other ordered chemical interfaces (*2*, *9*, *10*). Hydrogen bonding has often

Department of Chemistry, State University of New York, Stony Brook, NY 11794–3400, USA.

*To whom correspondence should be addressed. E-mail: nancy.goroff@sunysb.edu

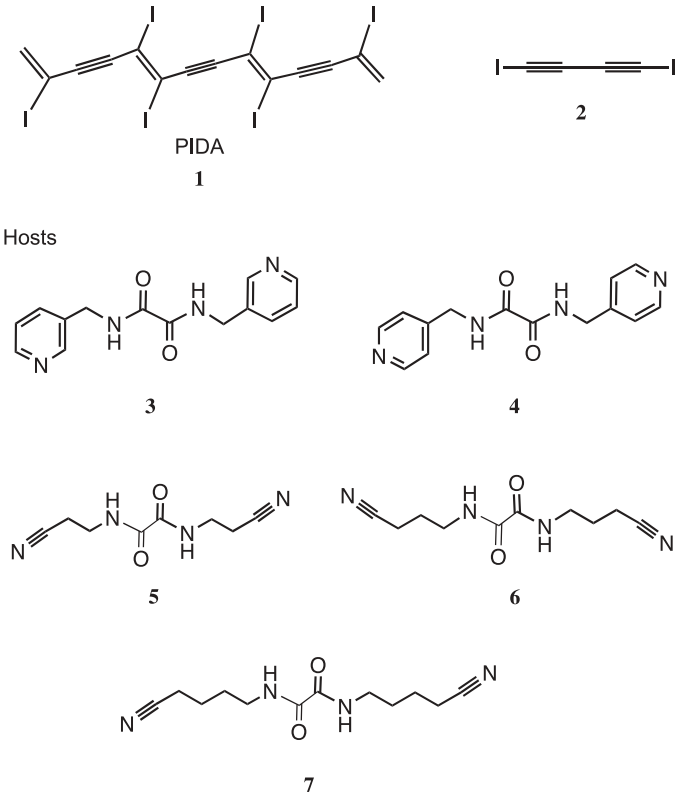


Fig. 1. PIDA (**1**) and monomer **2**, shown with several Lewis-basic host molecules, including new compounds **5** to **7**.

played a key role in these methods; for example, with monomers that do not align properly on their own, hydrogen bonding to a partner compound can produce co-crystals that have the appropriate repeat distance for 1,4-polymerization (*10*, *11*).

Nonetheless, the polydiacetylene structures accessible by these routes have been limited. For example, phenyl-substituted polydiacetylenes are desirable targets predicted to have particularly strong nonlinear optical absorptions, but there have been very few reported poly(aryldiacetylenes) (*12*, *13*). Monomers such as diphenylbutadiyne do not align properly for 1,4-polymerization on their own and do not contain hydrogen-bonding functional groups to interact with an appropriate co-crystallizing agent.

Poly(diiododiacetylene), or PIDA (**1**, Fig. 1), represents a possible solution to the general problem of polydiacetylene synthesis. PIDA contains the polydiacetylene backbone but with only halogen-atom substituents. The abundance of transition metal–catalyzed reactions available for adding substituents in place of a carbon-halogen bond makes PIDA a potential precursor to a wide variety of polydiacetylenes.

Polymer **1** may also act as a precursor to carbyne, the hypothesized linear *sp*-hybridized allotrope of carbon $(–C{\equiv}C–)_n$, which has thus far eluded definitive synthesis (*14*, *15*). Researchers have approached carbyne by rational synthesis of discrete polyyne rods (values of n from 2 to 14), allowing for the prediction of many of its properties (*16*–*19*). In addition, scientists have used electro- or thermochemical carbonization of small-molecule and polymeric substrates to make macroscopic carbyne-like materials consisting of polyyne domains and sp^2 hybridized kinks clustered around impurities in the samples. Hlavatý and co-workers have examined diiodobutadiyne (**2**), contained within the zeolite MCM-41, as a precursor to carbyne and found that ultraviolet irradiation could remove roughly half the iodine atoms of the starting material (*20*). However, producing a carbyne sample with a uniform structure that can be fully characterized remains a major challenge.

PIDA (**1**) is a promising precursor to fully ordered carbyne. With its linear arrangement of carbon atoms and small number of substituents, transformation of PIDA to carbyne requires only the cleaving of the carbon-iodine bonds and no further rearrangements. Electrochemical reduction of single-crystal PIDA samples would provide a previously unknown route to carbyne as an ordered material.

To prepare polymer **1** requires ordered alignment of the monomer diiodobutadiyne (**2**). Diiodobutadiyne was first reported by Baeyer in 1885 (*21*). It is moderately stable, decomposing rapidly and sometimes explosively when heated above its melting temperature of about 95°C (*21*–*23*). At room temperature, it can explode if subjected to physical shock, and it decomposes slowly under ambient light, forming an insoluble material with a graphite-like appearance (*22*). The decomposed material also explodes when heated, again releasing iodine gas (*21*). However, diiodobutadiyne can be handled safely at room temperature in gram quantities and decomposes very slowly if stored at –10°C in the dark. In solution, if exposed to light, it undergoes a very slow disproportionation reaction to give first tetraiodobutatriene and then hexaiodobutadiene, along with carbon-rich polymeric material (*20*, *24*).

Although compound **2** has been known for over a century, efforts to determine its crystal structure precisely have been unsuccessful. As

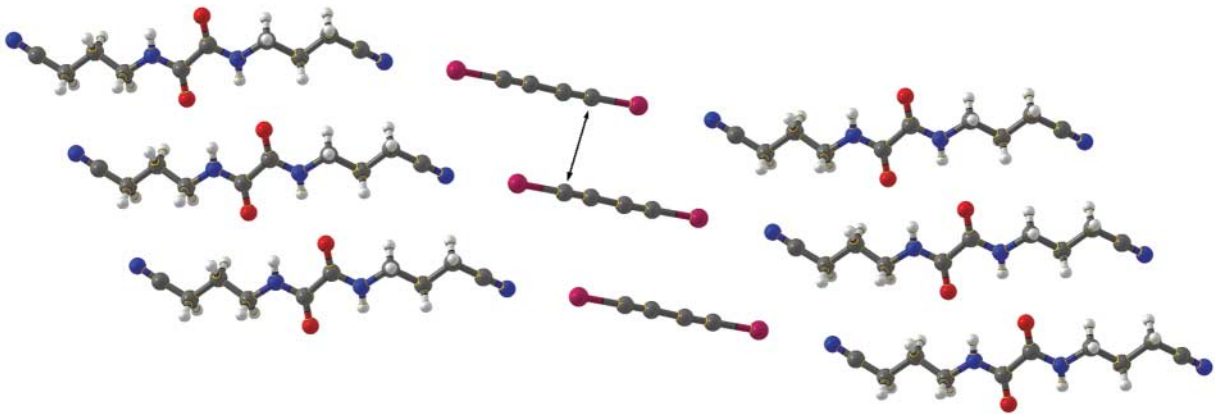


Fig. 2. Molecular packing in the co-crystal of **2** and **6**. The repeat distance is 5.25 Å, with a C1-C4′ intermolecular contact distance (indicated by the arrow) of 4.0 Å and a declination angle of 51.3°.

a linear, rod-shaped molecule, diyne **2** forms needle-shaped crystals that are disordered along the direction of the molecular axis. In these crystals, the alignment between adjacent monomers is random, preventing ordered polymerization. Thus, the observed decomposition products formed from solid **2** do not have the structure of polymer **1** but instead are disordered materials of varying composition and connectivity.

To create ordered crystals of diyne **2** with controlled molecular spacing, we take advantage of another property of the monomer, namely the strong Lewis acidity of its iodine substituents (*25*–*27*). Iodine atoms in general are highly polarizable and therefore Lewis acidic. In iodoalkynes, this effect is magnified because sp hybridization increases the electronegativity of the carbon atom bonded to iodine, polarizing the C-I bond toward carbon. Gas-phase calculations and solution-phase experiments indicate that the interactions between iodoalkynes and individual Lewis bases are somewhat weaker than comparable hydrogen bonds (*27*–*29*). Even so, such interactions between Lewis-acidic iodines and appropriate bases are often called "halogen bonds," in analogy to the more common hydrogen bond, and they have been demonstrated as a reliable organizational motif in the solid state (*30*–*32*). Iodoalkyne halogen bonds have been used by Yamamoto, Kato, and co-workers, who have prepared multicomponent materials containing organic radical cation salts and neutral iodoalkynes (*33*, *34*). In these systems, interactions between the Lewis-acidic iodoalkyne and the halide counter ion of the salt determine the spacing of the cations and therefore the electronic properties of the materials.

Alone, rod-shaped diiodobutadiyne forms crystals that are disordered along the long dimension of the molecule and therefore unsuitable for controlled topochemical polymerization. We previously demonstrated (*32*) that the Lewis acidity of the iodine atoms in **2** can be exploited to align the diyne in co-crystals with Lewis-basic hosts such as **3** and **4** (Fig. 1). However, efforts to polymerize **2** topochemically to form PIDA in these crystals have been unsuccessful, hampered by steric crowding around the iodines. To overcome these steric problems, we have now prepared a series of new host molecules, compounds **5** to **7** (*35*).

Like hosts **3** and **4**, compounds **5** to **7** each contain a central oxalamide group, designed to act as the primary organizing structure in the co-crystals. Fowler, Lauher, and co-workers have demonstrated that oxalamides form self-complementary hydrogen bonding networks with a repeat distance (4.9 to 5.1 Å) that matches well the 4.9 Å distance requirement for diyne 1,4-polymerization (*36*, *37*). In compounds **5** to **7**, linear alkyl nitrile chains replace the planar pyridyl groups of hosts **3** and **4**. We have prepared these new hosts from the corresponding aminoalkyl nitriles by condensation with diethyl oxalate. Although nitriles are weaker Lewis bases than pyridine nitrogen, previous experiments have indicated that they do act as electron donors to iodoalkynes in solution (*27*).

To form co-crystals of monomer **2** with each host, we prepared a 1:1 solution with 0.01 M concentration in methanol. Slow evaporation provided crystals for further analysis. From experiments with the three new hosts, we obtained three different results. Host **5** has much lower solubility than that of diyne **2**, and it precipitated from solution alone. However, both **6** and **7** formed co-crystals with **2**. In the case of host **6**, these co-crystals contained monomer **2**, as expected. In the case of host **7**, however, the isolated co-crystals contained PIDA, polymer **1**, which formed spontaneously under the crystallization conditions.

In the crystal structure of the **2•6** co-crystal (Fig. 2), much like the co-crystals of host **3** or **4** with **2**, the nitrile nitrogens are halogen-bonded to the iodines of **2**, but the repeat dis-

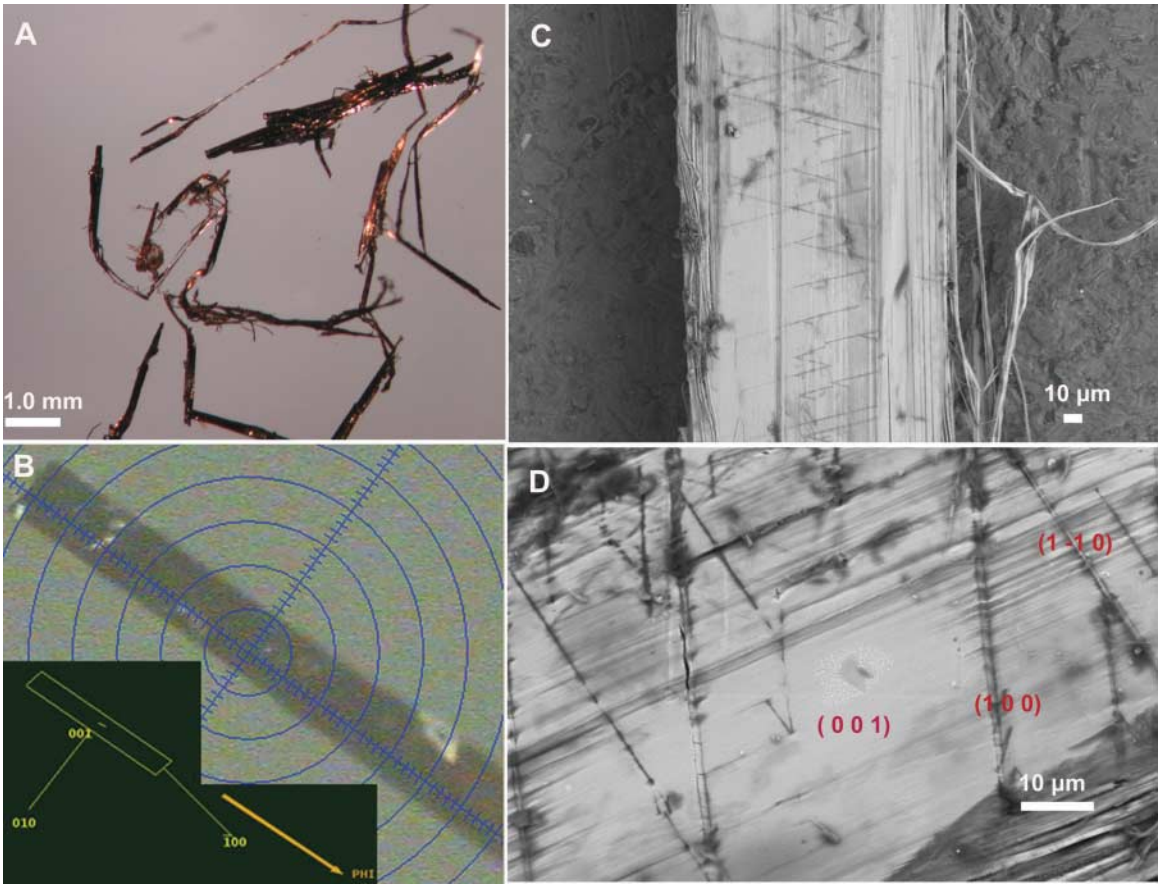


Fig. 3. Morphology of PIDA. (**A**) Microscopic image taken under polarized light with a Nikon SMZ800 optical microscope (Nikon Instruments, Incorporated, Melville, New York), (**B**) crystal faces identified by single-crystal x-ray diffraction analysis, and (**C** and **D**) field-emission SEM images of a crystalline PIDA fiber (FE-SEM LEO 1550, LEO Electron Microscopy, Thornwood, New York).

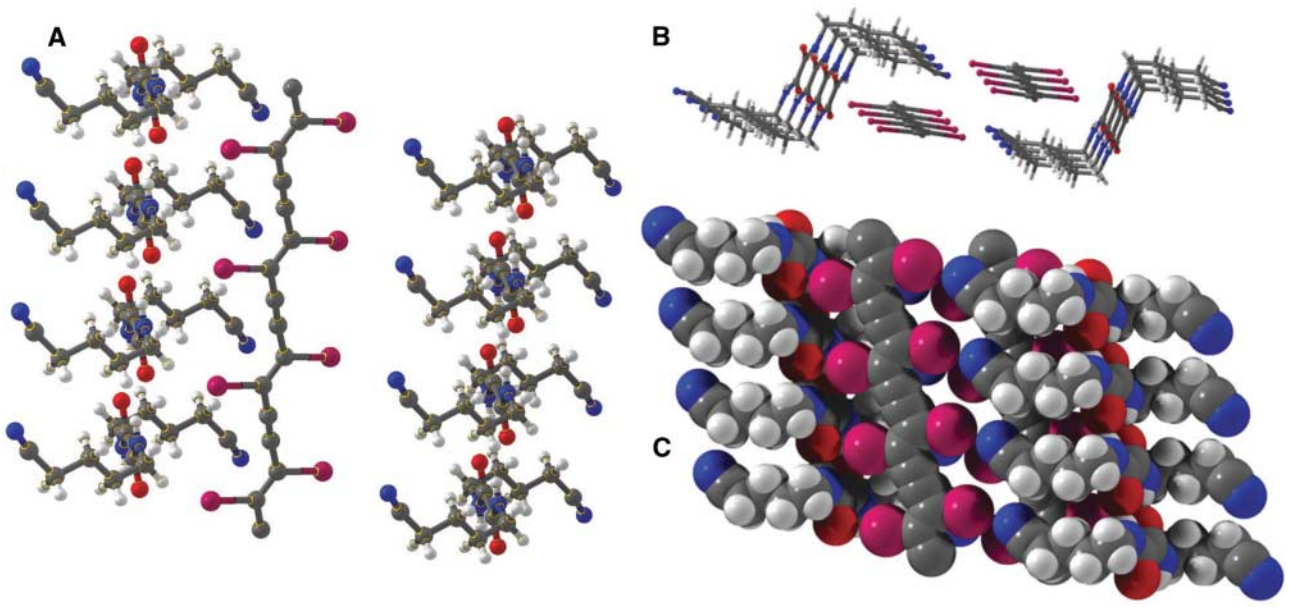


Fig. 4. Co-crystals of PIDA (**1**) and host **7**, as determined by x-ray diffraction (*39*). (**A**) A single polymer chain and hosts, (**B**) top view in ball-and-stick representation, and (**C**) side view in space-filling representation.

tance (5.25 Å) is longer than the targeted 4.9 Å needed for polymerization. Attempts to induce polymerization in these co-crystals have thus far been unsuccessful.

Whereas the **2•6** co-crystals are brown in color, the co-crystals isolated from solutions containing **2** and **7** are deep blue-purple; under reflected light, they have a coppery appearance and look like wire filaments (Fig. 3A). X-ray diffraction experiments established that these co-crystals contain polymer **1** rather than monomer **2** (Fig. 4). In the crystal structure, the polymer strands are all parallel to the oxalamide hydrogen-bonding network, with a repeat distance of 4.94 Å, close to the expected value of 4.9 Å. The host nitriles are each halogen-bonded to iodine atoms, as predicted. Unexpectedly, however, half the iodine atoms of **1** in the crystal have close contacts to oxalamide oxygen atoms rather than nitrile nitrogens, leading to a stoichiometry of 1:2 instead of the predicted 1:1 ratio. The oxygen atoms of host **7** are simultaneously acting as hydrogen-bond acceptors, suggesting that any stabilization from the oxygen-iodine interaction is likely quite small, although it contributes to the crystal packing structure.

We have been able to prepare co-crystals of **1** and **7** with lengths of up to 2 cm. Scanning electron microscopy (SEM) reveals the fibrous nature of these crystals, consistent with the alignment of the polymer strands within each crystal (Fig. 3, C and D). X-ray diffraction experiments confirm that the long dimension in these crystals corresponds to the (100) axis, the direction of polymer growth. In addition, the SEM images of the crystals include periodic linear steps, with heights of about 2 μm, either parallel to one another or crossing at an angle of 30°. X-ray diffraction experiments identify these steps as corresponding to the (100) and (110) faces of the crystal, with the large flat surface corresponding to the (001) face.

By using host **7**, we have not obtained co-crystals containing unreacted monomer **2**. However, comparison with the **2•6** co-crystal structure indicates that the atoms of the guest likely undergo substantial motion upon polymerization (*38*). The flexibility of the host and the relative weakness of the interactions between host and guest may lower the barrier to such motion, making the polymerization reaction especially favorable.

At the same time, this flexibility and lack of directional preference may also be the source of some experimental difficulties. Although host **7** and monomer **2** reliably form polymer **1** in crystallite films from methanol solution, single crystals large enough for individual analysis, such as described above, can only be obtained haphazardly. Nonetheless, our studies of the crystallites indicated that they also contain polymer **1**. The electronic absorption spectra from these samples (fig. S1) exhibited a double-peak absorption with maxima at 594 nm (minor) and 650 nm (major), characteristic of an ordered polydiacetylene in the "blue phase," where the polymer chains have fully planar, *s* trans conformation (*2*).

Although we have not yet measured the electrical conductivity of PIDA, the ability to obtain SEM images of the material at low voltages indicates qualitatively that polymer **1** is conducting. In addition, the SEM images highlight a benefit of topochemical synthesis. Because the polymer strands in the co-crystal are aligned, the electrical and optical properties of **1** will be highly anisotropic, with electronic and excitonic mobilities much higher along the (100) axis than in the other two dimensions. This anisotropy will allow for greater control in devices prepared from PIDA. At the same time, for solution-based applications, Lewis-basic solvents or co-solvents will interact with the iodine atoms of polymer **1**, increasing its solubility. The iodine atoms of **1** will therefore play a key role, not only in its synthesis, but also in future applications.

References and Notes

1. H. Shirakawa, E. J. Louis, A. G. MacDiarmid, C. K. Chiang, A. J. Heeger, *J. Chem. Soc. Chem. Comm.* **1977**, 578 (1977).
2. R. W. Carpick, D. Y. Sasaki, M. S. Marcus, M. A. Eriksson, A. R. Burns, *J. Phys. Cond. Matter* **16**, R679 (2004).
3. J. M. Nunzi, *C. R. Phys.* **3**, 523 (2002).
4. C. D. Dimitrakopoulos, P. R. L. Malenfant, *Adv. Mater.* **14**, 99 (2002).
5. D. T. McQuade, A. E. Pullen, T. M. Swager, *Chem. Rev.* **100**, 2537 (2000).
6. J. W. Y. Lam, B. Z. Tang, *Acct. Chem. Res.* **38**, 745 (2005).
7. G. Wegner, *Z. Naturforsch. Teil B* **24**, 824 (1969).
8. R. H. Baughman, *J. Polymer Sci. Polymer Phys. Ed.* **12**, 1511 (1974).
9. D. Bloor, R. R. Chance, Eds., *Polydiacetylenes: Synthesis, Structure and Electronic Properties*, no. 102 of *NATO ASI Series E: Applied Sciences* (Nijhoff, Dordrecht, 1985).
10. F. W. Fowler, J. W. Lauher, *J. Phys. Org. Chem.* **13**, 850 (2000).
11. J. J. Kane, R. F. Liao, J. W. Lauher, F. W. Fowler, *J. Am. Chem. Soc.* **117**, 12003 (1995).
12. A. Sarkar, S. Okada, H. Matsuzawa, H. Matsuda, H. Nakanishi, *J. Mater. Chem.* **10**, 819 (2000).
13. Y. H. Chan, J. T. Lin, I. W. P. Chen, C. H. Chen, *J. Phys. Chem. B* **109**, 19161 (2005).
14. L. Kavan, *Chem. Rev.* **97**, 3061 (1997).
15. R. B. Heimann, S. E. Evsyukov, L. Kavan, *Carbyne and Carbynoid Structures* (Kluwer Academic, Dordrecht, 1999).
16. S. Hino, Y. Okada, K. Iwasaki, M. Kijima, H. Shirakawa, in *Nanonetwork Materials: Fullerenes, Nanotubes, and Related Systems,* vol. 590 of *AIP Conference Proceedings* (Springer, New York, 2001), pp. 521–524.
17. T. Gibtner, F. Hampel, J.-P. Gisselbrecht, A. Hirsch, *Chem. Eur. J.* **8**, 408 (2002).
18. S. Eisler *et al.*, *J. Am. Chem. Soc.* **127**, 2666 (2005).
19. Q. Zheng, J. A. Gladysz, *J. Am. Chem. Soc.* **127**, 10508 (2005).

20. J. Hlavatý, J. Rathousky, A. Zukal, L. Kavan, *Carbon* **39**, 53 (2000).
21. A. Baeyer, *Ber. Dtsch. Chem. Ges.* **18**, 2269 (1885).
22. F. Straus, L. Kollek, H. Hauptmann, *Ber. Dtsch. Chem. Ges. B.* **63B**, 1886 (1930).
23. E. Heilbronner, V. Hornung, J. P. Maier, E. Kloster-Jensen, *J. Am. Chem. Soc.* **96**, 4252 (1974).
24. J. A. Webb, P.-H. Liu, O. L. Malkina, N. S. Goroff, *Angew. Chem. Int. Ed. Engl.* **41**, 3011 (2002).
25. C. Laurence, M. Queignec-Cabanetos, T. Dziembowska, R. Queignec, B. Wojtkowiak, *J. Am. Chem. Soc.* **103**, 2567 (1981).
26. C. Laurence, M. Queignec-Cabanetos, B. Wojtkowiak, *Can. J. Chem.* **61**, 135 (1983).
27. J. A. Webb, J. E. Klijn, P. A. Hill, J. L. Bennett, N. S. Goroff, *J. Org. Chem.* **69**, 660 (2004).
28. P. D. Rege, O. L. Malkina, N. S. Goroff, *J. Am. Chem. Soc.* **124**, 370 (2002).
29. W. N. Moss, N. S. Goroff, *J. Org. Chem.* **70**, 802 (2005).
30. P. Metrangolo, G. Resnati, *Chem. Eur. J.* **7**, 2511 (2001).
31. A. Crihfield *et al.*, *Cryst. Growth Des.* **3**, 313 (2003).
32. N. S. Goroff, S. M. Curtis, J. A. Webb, F. W. Fowler, J. W. Lauher, *Org. Lett.* **7**, 1891 (2005).
33. H. M. Yamamoto, J. I. Yamaura, R. Kato, *Synth. Met.* **102**, 1448 (1999).
34. H. M. Yamamoto, J. I. Yamaura, R. Kato, *Synth. Met.* **102**, 1515 (1999).
35. Materials and methods are available as supporting material on *Science* Online.
36. S. Coe *et al.*, *J. Am. Chem. Soc.* **119**, 86 (1997).
37. T. L. Nguyen, F. W. Fowler, J. W. Lauher, *J. Am. Chem. Soc.* **123**, 11057 (2001).
38. For example, in the **2•6** co-crystal, the iodine-nitrogen halogen bonding distance is 3.07 Å, and the C–I–N angle is 174.4°. In contrast, the iodine-nitrogen distance in the **1•7** co-crystal is 3.21 Å, and the C–I–N angle is 167.8°. The iodine-oxygen close contact is similar, with a distance of 3.20 Å and angle of 168.5°.
39. The unit cell parameters (triclinic, $P\bar{1}$) are as follows : a = 4.944(2) Å, b = 9.142(3) Å, c = 15.051(5) Å, α = 87.923(6)°, β = 84.590(6)°, γ = 78.759(7)°, V = 664.1(4) Å^3, Z = 1.

40. We thank N. S. Sampson and H. Tang for their help in measuring the electronic absorption spectrum of PIDA. We are grateful to NSF (grants CHE-9984937, CHE-0446749, and CHE-0453334) for financial support of this research. Crystallographic details for the **1•7** and **2•6** co-crystals are available free of charge from the Cambridge Crystallographic Data Centre under deposition numbers CCDC 294370 and CCDC 294371, respectively.

Supporting Online Material

www.sciencemag.org/cgi/content/full/312/5776/1030/DC1
Materials and Methods
Figs. S1 to S3
Tables S1 and S2

5 January 2006; accepted 17 March 2006
10.1126/science.1124621

Fast Mass Transport Through Sub–2-Nanometer Carbon Nanotubes

Jason K. Holt,[1]* Hyung Gyu Park,[1,2]* Yinmin Wang,[1] Michael Stadermann,[1] Alexander B. Artyukhin,[1] Costas P. Grigoropoulos,[2] Aleksandr Noy,[1] Olgica Bakajin[1]†

We report gas and water flow measurements through microfabricated membranes in which aligned carbon nanotubes with diameters of less than 2 nanometers serve as pores. The measured gas flow exceeds predictions of the Knudsen diffusion model by more than an order of magnitude. The measured water flow exceeds values calculated from continuum hydrodynamics models by more than three orders of magnitude and is comparable to flow rates extrapolated from molecular dynamics simulations. The gas and water permeabilities of these nanotube-based membranes are several orders of magnitude higher than those of commercial polycarbonate membranes, despite having pore sizes an order of magnitude smaller. These membranes enable fundamental studies of mass transport in confined environments, as well as more energy-efficient nanoscale filtration.

Carbon nanotubes, with diameters in the nanometer range and atomically smooth surfaces, offer a unique system for studying molecular transport and nanofluidics. Although the idea that water can occupy such confined hydrophobic channels is somewhat counterintuitive, experimental evidence has confirmed that water can indeed occupy these channels (*1*, *2*). Water transport through molecular-scale hydrophobic channels is also important because of the similarity of this system to transmembrane protein pores such as aquaporins (*3*). In recent years, numerous simulations (*4*, *5*) of water transport through single-walled carbon nanotubes (SWNTs) have suggested not only that water occupies these channels, but also that fast molecular transport takes place, far in excess of what continuum hydrodynamic theories would predict if applied on this length scale. Molecular dynamics (MD) simulations attribute this enhancement to the atomic smoothness of the nanotube surface and to molecular ordering phenomena that may occur on confined length scales in the 1- to 2-nm range (*4*, *5*). For similar reasons, simulations of gas transport through SWNTs (*6*) predict flux enhancements of several orders of magnitude relative to other similarly sized nanoporous materials. Membrane-based gas separations, such as those using zeolites (*7*), provide precise separation and size exclusion, although often at the expense of throughput or flux. It may be possible to use SWNTs to create a membrane that offers both high selectivity and high flux.

To investigate molecular transport on this length scale, we need to fabricate a carbon nanotube membrane that has a pore size of 1 to 2 nm. Researchers have recently fabricated multiwalled carbon nanotube (MWNT) membranes with larger pore diameters (6 to 7 nm) by encapsulation of vertically aligned arrays of MWNTs (*8*, *9*) and by templated growth within nanochannel alumina (*10*). Enhanced water transport through these larger MWNTs has recently been reported (*11*). Quantifying transport through an individual tube in a MWNT membrane is difficult, however, because MWNTs are prone to blockages, in particular by "bamboo" structures and catalyst particles that can migrate to and obstruct the nanotube interior (*9*, *12*, *13*). The consequence of such blockages is a marked reduction of the active membrane pore density. In contrast, there are few, if any, reports of "bamboo" structure formation or catalyst migration for SWNTs or double-walled carbon nanotubes (DWNTs). However, it is difficult to produce vertically aligned carbon nanotubes of this size (*14*, *15*). The major challenges also lie in finding a conformal deposition process to fill the gaps in this nanotube array, as well as in designing a selective etching process to open up the nanotube channels without producing voids in the membrane.

Table 1. Size exclusion tests on DWNT and MWNT membranes and molecular fluxes (per unit membrane area) of analytes. Values denoted by "<" were derived from the limits of detection for our concentration measurements when we did not observe any Au particles in the permeate solution. Differences of three to four orders of magnitude between this limiting value and the flux of the next smallest species indicate that the given analyte did not pass through the membrane.

Analyte	Analyte size (nm)	DWNT membrane flux (molecules cm^{-2} s^{-1})	MWNT membrane flux (molecules cm^{-2} s^{-1})
$Ru^{2+}(bipyr)_3$	1.3	5×10^{13}	5×10^{13}
Colloidal Au 1	2 ± 0.4	$<2 \times 10^{9}$	1×10^{11}
Colloidal Au 2	5 ± 0.75	$<3 \times 10^{8}$	3×10^{10}
Colloidal Au 3	10 ± 1	Not tested	$<4 \times 10^{7}$

[1]Chemistry and Materials Science Directorate, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA. [2]Department of Mechanical Engineering, University of California, Berkeley, CA 94720, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: bakajin1@llnl.gov

We have developed a microelectromechanical systems (MEMS)–compatible fabrication process (Fig. 1A) for sub–2-nm nanotube pore membranes. The process uses catalytic chemical vapor deposition to grow a dense, vertically aligned array of DWNTs on the surface of a silicon chip (Fig. 1B), followed by conformal encapsulation of the nanotubes by a hard, low-pressure chemical vapor–deposited silicon nitride (Si_3N_4) matrix (Fig. 1C). Scanning electron microscopy (SEM) images show that our process produces gap-free membranes over the length scale of the whole chip. The excess silicon nitride is removed from both sides of the membrane by ion milling, and the ends of the nanotubes are opened up with reactive ion etching. The membranes remain impermeable to both liquids and gases until the very last etching step; this is further evidence that our fabrication process produces crack-free and void-free membranes. Transmission electron microscopy (TEM) plan-view images (Fig. 2, C to E) of a slice of the membrane also demonstrate that the silicon nitride coats the DWNTs conformally and does not leave any gaps between the outer surface of the nanotube and the silicon nitride. We also used the same nitride-encapsulation method to produce MWNT membranes (*16*).

To characterize the membrane pore size, we performed size exclusion measurements (Table 1) (*16*) and compared these results with electron microscopy (EM) data. DWNT membranes passed species with sizes up to 1.3 nm yet blocked 2-nm gold particles, which suggests that these membranes have pore sizes between 1.3 and 2 nm. Comparison of the water flow rates before and during filtration, coupled with the upper-limit estimate of the flux of 2-nm gold particles, suggests that less than 0.1% of the flux through the membrane can be attributed to pores larger than 2 nm (*16*). These size exclusion measurements are further supported by the evidence obtained through EM. The distribution of DWNTs, as measured by TEM, revealed an inner diameter average of 1.6 nm (Fig. 2B). TEM images (Fig. 2C) of the membrane also revealed that the only holes that span the thickness of the membrane are of a size that is consistent with the inner diameter distribution of as-grown DWNTs. A MWNT membrane, used as a reference, transported colloidal gold particles with diameters of 2 and 5 nm but excluded the 10-nm colloidal gold particles, in agreement with MWNT diameters of 6.5 ± 0.5 nm estimated by TEM. This result also suggests that pore clogging by particles smaller than the average pore size is unlikely for the solution concentrations used in these experiments. We conclude that the transport in our samples occurs exclusively through the inner pores of the carbon nanotubes spanning the membrane.

The absolute gas flux through our membranes exceeded the flux predicted by the Knudsen diffusion model. As the dimensions

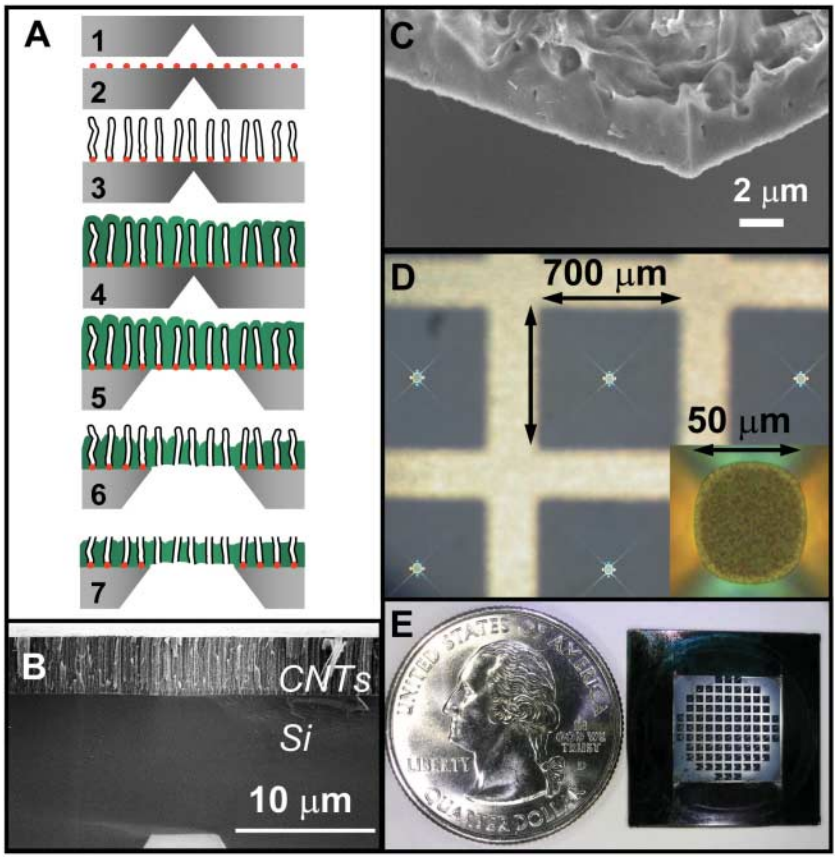


Fig. 1. (**A**) Schematic of the fabrication process. Step 1: microscale pit formation (by KOH etching). Step 2: catalyst deposition/annealing. Step 3: nanotube growth. Step 4: gap filling with low-pressure chemical vapor–deposited Si_3N_4. Step 5: membrane area definition (by XeF_2 isotropic Si etching). Step 6: silicon nitride etch to expose nanotubes and remove catalyst nanoparticles (by Ar ion milling); the membrane is still impermeable at this step. Step 7: nanotube uncapping (reactive ion etching); the membrane begins to exhibit gas permeability at this step. (**B**) SEM cross section of the as-grown DWNTs (CNTs). (**C**) SEM cross section of the membrane, illustrating the excellent gap filling by silicon nitride. (**D**) Photograph of the open membrane areas; inset shows a close-up of one membrane. (**E**) Photograph of the membrane chip that contains 89 open windows; each window is 50 μm in diameter.

Table 2. Comparisons of experimental air flow rates observed for several DWNT membranes with Knudsen model predictions, and of experimental water flow rates with continuum flow model predictions. The differences among the three DWNT membranes are most likely the result of different numbers of pores opened in the fabrication process. Values for a polycarbonate membrane are provided as a reference. Pore diameters were determined from size exclusion measurements, TEM measurements, and (for polycarbonate) manufacturer's specifications. Pore density values are upper limits, as determined from TEM measurements and (for polycarbonate) manufacturer's specifications.

Membrane	Pore diameter (nm)	Pore density (cm^{-2})	Thickness (μm)	Enhancement over Knudsen model* (minimum)	Enhancement over no-slip, hydrodynamic flow† (minimum)	Calculated minimum slip length‡ (nm)
DWNT 1	1.3 to 2.0	$\leq 0.25 \times 10^{12}$	2.0	40 to 120	1500 to 8400	380 to 1400
DWNT 2	1.3 to 2.0	$\leq 0.25 \times 10^{12}$	3.0	20 to 80	680 to 3800	170 to 600
DWNT 3	1.3 to 2.0	$\leq 0.25 \times 10^{12}$	2.8	16 to 60	560 to 3100	140 to 500
Polycarbonate	15	6×10^{8}	6.0	2.1	3.7	5.1

*From (*18*). †From (*26*). ‡From (*29*).

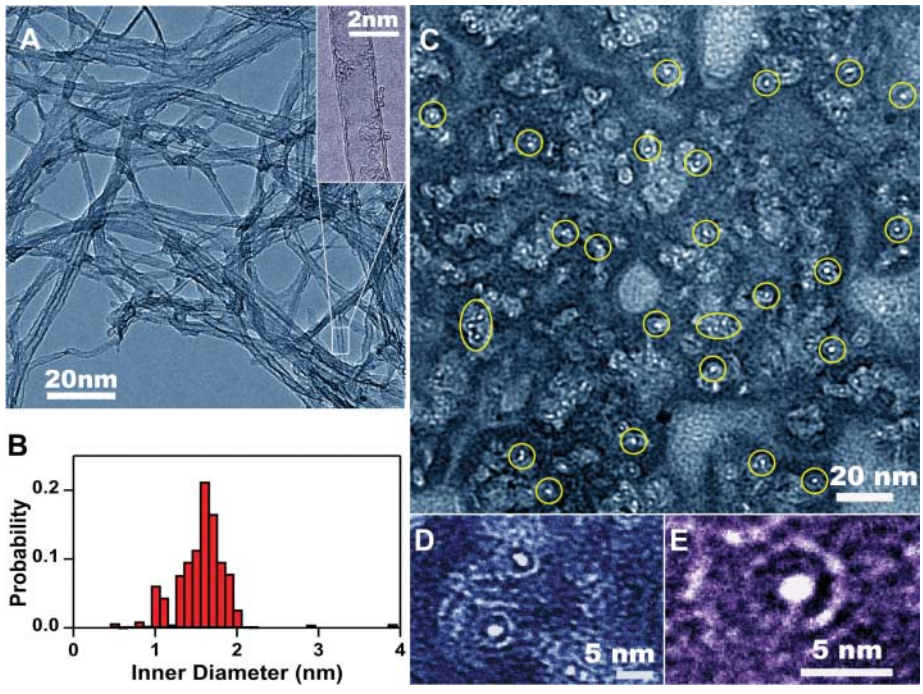


Fig. 2. (**A**) TEM images of as-grown nanotubes, prepared by removing them from the silicon substrate and dispersing them in dimethylformamide. The majority of the carbon nanotubes are double-walled, as identified in the high-resolution inset. (**B**) Pore size distribution, derived from TEM measurements of the inner diameter of 391 individual carbon nanotubes, reveals an average pore size of 1.6 nm. The average outer diameter of these DWNTs is estimated to be 2.3 nm. (**C** to **E**) Plan-view TEM images of carbon nanotube membrane taken with the beam parallel to the nanotube axis. In (C), the nanotube membrane shows continuous nitride coating on the scales examined in this image (~0.2 μm by 0.2 μm). No microcracks or microvoids can be seen. The bright white spots (circled in yellow) are carbon nanotube pores, which can be identified by the surrounding ring-shaped coating of silicon nitride. For clarity, not all visible nanotubes were circled. The density of carbon nanotubes is measured to be $\sim 2.5 \times 10^{11}$ cm^{-2} from several similar TEM images. In (D) and (E), high-resolution TEM images of selected areas from (C) show conformal coating of silicon nitride. The bright white spots in the images have the same inner diameter as the carbon nanotubes.

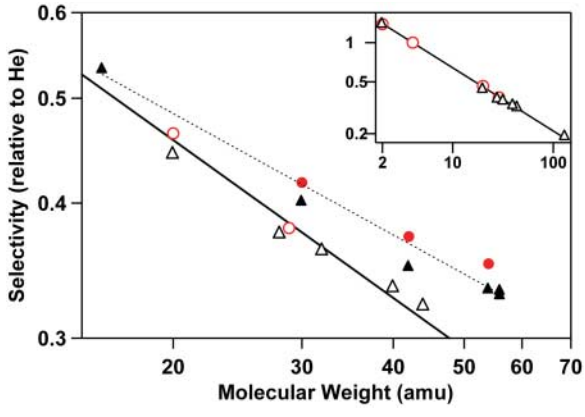


Fig. 3. Gas selectivity (defined as permeability relative to He) data for sub–2-nm DWNT (triangles) and MWNT (circles) membranes. Open symbols denote nonhydrocarbon gases (H_2, He, Ne, N_2, O_2, Ar, CO_2, Xe); solid symbols denote hydrocarbon gases (CH_4, C_2H_6, C_3H_6, C_4H_6, C_4H_8). The solid line is a power-law fit of the nonhydrocarbon gas selectivity data, showing a scaling predicted by the Knudsen diffusion model (exponent of –0.49 ± 0.01). The dashed line is a power-law fit of the hydrocarbon gas data, showing a deviation from the Knudsen model (exponent of –0.37 ± 0.02). The inset shows the full mass range of the nonhydrocarbon gas data, again illustrating agreement with the Knudsen model scaling.

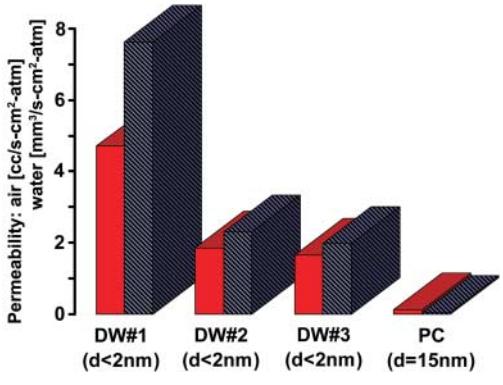


Fig. 4. Air (red) and water (blue) permeability as measured for three DWNT membranes (DW#1, 2, and 3) and a polycarbonate membrane (PC). Despite considerably smaller pore sizes, the permeabilities for all DWNT membranes greatly exceed those of the polycarbonate membrane.

of the pore shrink, the mean free path (λ) becomes larger than the channel dimensions (*d*) and the transport enters the molecular flow regime. In such situations, where particle-surface collisions dominate over particle-particle collisions, the Knudsen diffusion model (*17*) is frequently applied. Indeed, our pore geometries are characterized by Knudsen numbers (λ/*d*) of 10 to 70, which places them well into the free molecular transport regime. However, the flux measured through our membranes exceeded the flux predicted by the Knudsen model (*18*) by at least one to two orders of magnitude (Table 2). By comparison, a polycarbonate membrane (Nucleopore, Osmonics Inc.) revealed just a slight enhancement in flux. The single largest uncertainty in quantifying the flux through our membrane pores lies in determination of the active pore density (i.e., those nanotubes that are open and spanning the membrane). A pore density estimate of 2.5×10^{11} cm^{-2} was derived from the plan-view TEM images of the DWNT membrane (Fig. 2, C to E), and the enhancement factors that we report represent lower boundary estimates (*19*). The observed flow enhancement is most likely caused by the intrinsic smoothness of the nanotube surface, as predicted by MD simulations of gas flow through SWNTs (*6*, *20*–*22*). In atomically smooth pores, the nature of gas-wall collisions can change from purely diffuse (as in the Knudsen model) to a combination of specular and diffuse collisions (*23*), thus leading to observed faster transport.

We found that single-component selectivity for most of the gases exhibited the expected inverse-square-root scaling of molecular mass (Fig. 3, inset) with the exception of hydrocarbons, whose selectivities were higher. This result is not surprising for a molecular diffusion process because it reflects the mass dependence of molecular velocity (note that the nature of wall collisions would not affect the mass scaling). Of all the measured gases, only the hydrocarbons deviated from this pattern, exhibiting a higher selectivity (Fig. 3) in both DWNT

and MWNT membranes. Interestingly, a reference polycarbonate membrane with a pore size of 15 nm did not show this deviation. We attribute the deviation to the preferential interaction of hydrocarbons with the carbon nanotube sidewalls. The hydrocarbon transport enhancement most likely results from surface diffusion or possibly a solubility/diffusion mechanism (*24*). Pulse mass analysis of various organic compounds has shown strong adsorption of hydrocarbon molecules (e.g., hexane) on SWNTs relative to more polar molecules (e.g., ethanol) (*25*). It is noteworthy that the hydrocarbon selectivity we observe in these single-component experiments may be more pronounced for practical gas separation problems where mixtures are involved (*20*).

Our membranes also transported water across the carbon nanotube channels at rates that cannot be accounted for by continuum flow models. The measured water flow rates reveal a flow enhancement (Table 2) that is more than three orders of magnitude faster than the no-slip, hydrodynamic flow as calculated from the Hagen-Poiseuille equation (*26*). Breakdown of this continuum model is not surprising for channels 1 to 2 nm in size. If we take the formalism used for gases and define a mean free path in liquids as the molecular diameter (e.g., ~0.3 nm for H_2O), the Knudsen number for a 1- to 2-nm pore is 0.15 to 0.3. These values lie on the border between "slip flow" and "transitional flow." In this size regime, where the pore is only ~7 water molecules in diameter, continuum theory concepts such as a velocity profile may be difficult to define. For this reason, MD simulations are often used for the prediction of water flows through carbon nanotube pores with diameters on the order of 1 nm (*4*, *5*). However, the computational expense of MD simulations, as well as observations of a finite fluid "slipping velocity" at hydrophobic interfaces (*27*), have motivated attempts to use meso- and macroscopic flow models to simulate flow through SWNTs (*28*). These simulations calculated a corresponding "slip length" that describes the noncontinuum behavior of a liquid near the pore walls. If we apply a similar formalism for the flow through our sub–2-nm nanotube membranes, we calculate (*29*) slip lengths as large as 1400 nm (Table 2). These values are almost three orders of magnitude larger than the pore size and are on the order of the overall size of the system (pore length). In contrast, the polycarbonate membrane with a pore size of 15 nm reveals a much smaller slip length of just 5 nm. This comparison suggests that slip-flow formalism may not be applicable to water flow through sub–2-nm carbon nanotubes, possibly because of length scale confinement (*30*) or partial wetting of the carbon nanotube surface (*31*).

Our observed water flux compares well with that predicted by the MD simulations (*5*). The simulations predict a flux of 12 water molecules through 1 nm^2 of the nanotube cross-sectional area in 1 ns; our measured flux, extrapolated to the simulation pressure drop, corresponds to 10 to 40 water molecules nm^{-2} ns^{-1} (*32*). The MD simulations attributed the observed high water flow rates to the formation of water "wires" in the confined space inside the nanotube. The strong dependence of the structure of the water in the nanotube on diameter (*33*) indicates that small differences in nanotube diameter can have large effects on transport. Therefore, it is unclear whether the mechanism proposed by MD is responsible for the high water flow rates observed with the larger nanotubes used in our experiments, or whether the flow enhancement can be attributed simply to the presence of a nearly frictionless surface.

Membrane permeability provides a figure of merit for membrane performance for practical applications. Despite having an order of magnitude smaller pore size, the enhanced flow rate per pore and the higher pore density render the sub–2-nm membranes superior to conventional polycarbonate membranes in both air and water permeability (*34*) (Fig. 4).

References and Notes

1. A. I. Kolesnikov *et al.*, *Phys. Rev. Lett.* **93**, 035503 (2004).
2. N. Naguib *et al.*, *Nano Lett.* **4**, 2237 (2004).
3. P. Agre *et al.*, in *Aquaporins*, S. Hohmann, S. Nielsen, P. Agre, Eds., vol. 51 of *Current Topics in Membranes* (Academic Press, San Diego, CA, 2001), pp. 1–38.
4. G. Hummer, J. C. Rasaiah, J. P. Noworyta, *Nature* **414**, 188 (2001).
5. A. Kalra, S. Garde, G. Hummer, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 10175 (2003).
6. A. I. Skoulidas, D. M. Ackerman, J. K. Johnson, D. S. Sholl, *Phys. Rev. Lett.* **89**, 185901 (2002).
7. Z. Lai *et al.*, *Science* **300**, 456 (2003).
8. B. J. Hinds *et al.*, *Science* **303**, 62 (2004); published online 26 November 2003 (10.1126/science.1092048).
9. J. K. Holt, A. Noy, T. Huser, D. Eaglesham, O. Bakajin, *Nano Lett.* **4**, 2245 (2004).
10. J. Li, C. Papadopoulos, J. M. Xu, M. Moskovits, *Appl. Phys. Lett.* **75**, 367 (1999).
11. M. Majumder, N. Chopra, R. Andrews, B. J. Hinds, *Nature* **438**, 44 (2005).
12. Y. Y. Wang, S. Gupta, R. J. Nemanich, Z. J. Liu, L. C. Qin, *J. Appl. Phys.* **98**, 014312 (2005).
13. H. Cui, O. Zhou, B. R. Stoner, *J. Appl. Phys.* **88**, 6072 (2000).
14. S. Maruyama, E. Einarsson, Y. Murakami, T. Edamura, *Chem. Phys. Lett.* **403**, 320 (2005).
15. K. Hata *et al.*, *Science* **306**, 1362 (2004).
16. See supporting material on *Science* Online.
17. A. F. Mills, *Mass Transfer* (Prentice-Hall, Upper Saddle River, NJ, 2001), pp. 68–69.
18. For gas flow in the Knudsen regime, the overall membrane flow rate can be determined from

$$Q_{gas} = \frac{2}{3}\sqrt{\frac{8\pi}{MRT}}(d/2)^3 V_m \frac{\Delta p}{L}\sigma A$$

where M is molecular weight, R is the universal gas constant, T is absolute temperature, d is pore diameter, V_m is the molar volume, Δp is the pressure drop, L is the thickness of the membrane, σ is the areal pore density, and $A = 89 \cdot \pi(25\ \mu m)^2 = 175{,}000\ \mu m^2$ is the total area of the membrane.
19. This density is only a factor of 4 smaller than the catalyst density on the substrate (~10^{12} cm^{-2}), also determined by TEM. This areal density is comparable to the measured areal density of SWNTs/DWNTs produced using a catalyst recipe similar to the one we used (*12*). The estimate from the TEM images still represents the upper bound for the density because it assumes that every DWNT that spans the section imaged in the TEM (thickness 50 nm) also spans the entire membrane thickness.
20. H. B. Chen, D. S. Sholl, *J. Membr. Sci.* **269**, 152 (2006).
21. D. M. Ackerman, A. I. Skoulidas, D. S. Sholl, J. K. Johnson, *Mol. Sim.* **29**, 677 (2003).
22. H. B. Chen, J. K. Johnson, D. S. Sholl, *J. Phys. Chem. B* **110**, 1971 (2006).
23. S. K. Bhatia, H. B. Chen, D. S. Sholl, *Mol. Sim.* **31**, 643 (2005).
24. C. Leger, H. D. Lira, R. Paterson, *J. Membr. Sci.* **120**, 187 (1996).
25. E. W. Bittner, M. R. Smith, B. C. Bockrath, *Carbon* **41**, 1231 (2003).
26. The Hagen-Poiseuille equation is

$$Q_{HP} = \frac{\pi(d/2)^4}{8\mu}\frac{\Delta p}{L}$$

where Q_{HP} is the volumetric flow rate, Δp is the pressure drop, d is the pore diameter, μ is the water viscosity, and L is the membrane thickness.
27. J. Baudry, E. Charlaix, A. Tonck, D. Mazuyer, *Langmuir* **17**, 5232 (2001).
28. E. M. Kotsalis, J. H. Walther, P. Koumoutsakos, *Int. J. Multiphase Flow* **30**, 995 (2004).
29. With the inclusion of a slip-flow correction, the Hagen-Poiseuille equation becomes

$$Q_{SLIP} = \frac{\pi\left[(d/2)^4 + 4(d/2)^3 L_s\right]}{8\mu}\frac{\Delta p}{L}$$

where L_s is defined as the slip length,

$$L_s = \frac{U_{wall}}{dU/dr}$$

where U_{wall} is the axial velocity at the wall, and dU/dr is the radial velocity gradient at the wall (or shear rate).
30. C. Cottin-Bizonne *et al.*, *Eur. Phys. J. E* **9**, 47 (2002).
31. V. S. J. Craig, C. Neto, D. R. M. Williams, *Phys. Rev. Lett.* **87**, 054504 (2001).
32. The simulation considered water transport across the carbon nanotubes driven by an osmotic pressure of about 100 atm. Our experiments used pressure drops of 1 atm. We have also observed linear dependence between the applied pressure drop and the flow rate across the membranes. As an approximation, we therefore used a linear extrapolation to compare our measured flows to the simulation results. We note two key differences between our experiments and the simulations: (i) The simulations used nanotubes 0.8 nm in diameter, whereas our samples were 1.6 nm in diameter on average; and (ii) the pressure drops were ~100 atm in the simulations versus 1 atm in our experiments, and it is unclear whether our linear extrapolation in flow rate versus pressure drop is valid over this range.
33. K. Koga, G. T. Gao, H. Tanaka, X. C. Zeng, *Nature* **412**, 802 (2001).
34. Permeability is defined as the volumetric flux, normalized by the pressure drop.

35. We thank J. Muyco for help with Raman spectroscopy, R. Friddle for assistance with atomic force microscopy measurements, W. J. Moberlychan for his assistance on FIB and TEM experiments, and D. Eaglesham for early contributions to the project and lively discussions. This work was performed under the auspices of the U.S. Department of Energy by University of California Lawrence Livermore National Laboratory under contract W-7405-Eng-48 with funding from the Laboratory Directed Research and Development Program. H.G.P. and A.B.A. were supported by a Student Employee Graduate Research Fellowship.

Supporting Online Material

www.sciencemag.org/cgi/content/full/312/5776/1034/DC1
Materials and Methods
Fig. S1

15 February 2006; accepted 15 March 2006
10.1126/science.1126298

Apes Save Tools for Future Use

Nicholas J. Mulcahy and Josep Call*

Planning for future needs, not just current ones, is one of the most formidable human cognitive achievements. Whether this skill is a uniquely human adaptation is a controversial issue. In a study we conducted, bonobos and orangutans selected, transported, and saved appropriate tools above baseline levels to use them 1 hour later (experiment 1). Experiment 2 extended these results to a 14-hour delay between collecting and using the tools. Experiment 3 showed that seeing the apparatus during tool selection was not necessary to succeed. These findings suggest that the precursor skills for planning for the future evolved in great apes before 14 million years ago, when all extant great ape species shared a common ancestor.

Tulving (*1*) recounts an Estonian tale of a girl who dreamed about attending a party but was unable to eat her favorite dessert because there were no spoons available. Facing the possibility of attending the party again, she took a spoon to bed. Crucially, the girl took the spoon not because she currently needed it, but because she would need it in the future. Tulving used this example to illustrate the putatively unique human ability to think about the past and plan for the future (*2*–*4*) and proposed that an analogous "spoon" test could be used to test for future planning in nonhuman animals. Future planning is cognitively demanding because it imposes a long delay between performing an action and getting rewarded for it: a skill that humans use when preparing a suitcase before a trip or by making a cake to celebrate someone's birthday. Although various animals can plan and execute multiple actions toward a goal (*5*, *6*), they may achieve this without taking into account future needs, just current ones (*3*, *4*, *7*). Thus, when chimpanzees transport stones to use them to crack open nuts, or New Caledonian crows make hook-shaped tools to fish for insects, they do so in an attempt to satisfy their current hunger state, not some future one.

Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, D-04103 Leipzig, Germany.

*To whom correspondence should be addressed. E-mail: call@eva.mpg.de

To date, the only evidence suggesting future planning in animals is the cache protection strategy of scrub jays (*8*). These birds transport food from old sites and re-cache it in new ones to avoid losing their caches to conspecifics that observed the initial location of the cache sites, thus saving the items for future consumption. Some researchers argue that future planning may have evolved independently in various taxa (*9*), whereas others argue that the nonhuman data do not truly capture the essence and complexity of human mental time travel (*1*, *4*). For instance, it is unknown whether animals would also transport and save nonedible items that would enable them to get food at a later time, an ability that may have played a crucial role in human evolution (*10*–*12*). Would extant apes also transport tools that they do not currently need, anticipating that they will need them in the future? This would be an important addition to the complex behavior displayed by scrub jays, because what is saved is not the food itself but a means to get the food, which is one step removed from the goal itself. The tool has no value in itself; it has value only in relation to the food.

We investigated future planning in bonobos and orangutans, because they represent our closest and most distant great ape relatives, respectively. This allowed us to make inferences about the possible time of emergence of the precursor skills for future planning in this group. If both species showed this skill, this would suggest that it may have evolved before 14 million years ago (Ma), when all apes shared a common ancestor (*13*); whereas if only bonobos showed it, it may have evolved within the past 14 My. Its absence in both species would suggest that future planning is a human adaptation that appeared within the past 7 My.

We tested five bonobos and five orangutans (table S1). First, subjects learned to use a tool to get a reward from an apparatus in the test room (*14*). Then, we placed two suitable and six unsuitable tools in the test room but blocked subjects' access to the baited apparatus. After 5 min, subjects were ushered outside the test room into the waiting room, and the caretaker removed all objects left in the test room while subjects watched. One hour later, subjects were allowed to return to the test room and were given access to the apparatus. Thus, to solve the problem, subjects had to select a suitable tool from the test room, bring it into the waiting room, keep it there for 1 hour, and bring it back into the test room upon their return (fig. S1). The trial ended after the subject retrieved the reward or 5 min had elapsed.

Subjects solved the problem an average of seven times (SEM = 1.8), with all subjects succeeding at least once within the first seven trials (Table 1). Subjects performed the key behavior of transporting tools out of the test room in 70% of the trials, targeting suitable tools significantly more often than would be expected by chance (t_5 = 2.59, P = 0.049) (expected = 25%, observed = 40.9%). Moreover, they returned to the test room 77.5% of the suitable tools but only 32.6.5% of the unsuitable tools (t_5 = 2.73, P = 0.041) (Fig. 1). Two orangutans (Dokana and Toba) on one and three occasions, respectively, brought back an inappropriate tool but broke off some small piece from it with which they still obtained the reward. Even if those trials are excluded from the analyses, subjects still initially targeted suitable tools (t_5 = 3.09, P = 0.027) and brought them back more often than

Table 1. Number and order of correct trials for each subject for each experiment. NT, not tested.

Subjects	Experiment 1		Experiment 2		Experiment 3		Experiment 4	
	Correct trials (*n*)	Trial no.	Correct trials (*n*)	Trial no.	Correct trials (*n*)	Trial no.	Correct trials (*n*)	Trial no.
Bonobos								
Kuno	7/16	7, 8, 10, 13–16	8/12	2–6, 9–11	7/16	5, 7, 8, 10, 11, 13, 15	NT	
Joey	2/16	1, 13	NT		NT		NT	
Limbuko	5/16	7, 10, 11, 13, 16	NT		6/16	5–7, 13, 14, 16	NT	
Yasa	NT		NT		NT		0/16	
Ulindi	NT		NT		NT		0/16	
Orangutans								
Walter	6/16	4–8, 14	NT		NT		NT	
Toba	7/16	1, 3, 6, 10–13	NT		6/16	1, 4, 6, 7, 12, 13	NT	
Dokana	15/16	1, 3–16	7/12	2, 4–6, 9, 11, 12	7/16	6–9, 11, 14, 16	NT	
Dunja	NT		NT		NT		2/16	2, 16
Pini	NT		NT		NT		5/16	1, 13–16

unsuitable tools (t_5 = 2.77, P = 0.04). Tool transport did not occur simply because subjects were already holding the tool when the door was opened after the delay period. Indeed, the opposite was true. On average, subjects were already holding the tool in only 8.2% (SEM = 5.3) of the successful trials when the door opened to return to the test room (this figure was 26.5% for the tools taken out of the test room). If those trials are excluded, subjects still solved the task on an average of 4.7 trials (out of 16).

In experiment 2, we increased the delay between tool retrieval and reward retrieval to 14 hours. We brought the subject into the test room, then sent her to the sleeping room (which served as waiting room) for the night, and brought her back to the test room in the morning. The sleeping room was not contiguous to the test room but located one floor above it. We tested one orangutan and one bonobo. Neither ape took any tools in the first trial. The orangutan took suitable tools in all 11 remaining trials (binomial test: P = 0.01, one-tailed), which she brought back and used in 7 trials to get the reward. The bonobo took suitable tools in 8 of the remaining 11 trials (binomial test: $P < 0.001$, one-tailed), which he always brought back to the testing room to get the reward.

In experiment 3, two bonobos and two orangutans had to use a hook to get an out-of-reach juice bottle suspended from a string (fig. S1). After subjects had learned to use the hook appropriately, they were presented with the hook and three unsuitable tools in the absence of the apparatus (or the reward). Five minutes later, subjects were ushered outside of the test room and brought back 1 hour later. Once subjects were inside the test room, the apparatus was installed and subjects could get the reward if they had returned with the hook. This experiment addressed two outstanding issues. It tested whether subjects would select the suitable tool in the absence of the apparatus or the reward. It controlled for the possibility that subjects took the tool in previous experiments to reduce their current hunger state, because the hook had been associated only with liquid procurement. Moreover, because water was available ad libitum, it is unlikely that subjects took the hook to reduce their thirst. Subjects' performance was comparable to that in experiment 1. Subjects solved the task on an average of 6.5 trials (SEM = 0.3), succeeding at least once within the first six trials, and transporting suitable tools more often than would be expected by chance (t_3 = 2.69, P = 0.037 one-tailed) (expected = 25%, observed = 41.5%). Moreover, they returned more suitable than unsuitable tools (79.6% versus 47.2%), but the difference was not statistically significant (t_3 = 2.25, P = 0.055, one-tailed) (Fig. 1).

Experiment 4 established the baseline probability of transporting tools in the absence of a future task but using identical reinforcement contingencies as in experiment 3. Two bonobos and two orangutans received the same treatment as in experiment 3, except that no apparatus was set up upon their return to the test room although they were rewarded if they brought the suitable tool back. Subjects solved the task significantly less often (mean = 1.8, SEM = 1.2) than did those in experiment 3 (t_6 = 3.91, P = 0.008). In fact, only two of the four subjects brought back the suitable tool at all, and they behaved differently from other successful subjects because after their first successful trial, they failed the next 11 and 14 trials, respectively (Table 1). Subjects in experiment 4 also solved the task significantly less often than those in experiment 1 (t_8 = 2.81, P = 0.023), thus ruling out the possible confounding effects of practice, because both groups of subjects were naïve when their respective experiments began.

Apes selected, transported, and saved a suitable tool not because they currently needed it but because they would need it in the future. These data taken all together cannot be easily explained by invoking traditional operant learning, because there was at least a 1-hour delay between the response (tool selection and transport) and the reinforcement. Typically, instrumental conditioning fails to occur if the response-reinforcement interval is greater than a few seconds in the absence of conditioned reinforcers (*3*, *15*, *16*), as evidenced by the poor performance of the subjects in experiment 4 as compared to other experiments. Arguably, the suitable tool could have acted as a conditioned reinforcer. However, this still does not explain the difference between experiment 4 and the other three experiments (fig. S2), which is even more remarkable because subjects were reinforced for returning with the tool in all experiments. Perhaps subjects did not transport the tool in experiment 4 because they had not been reinforced to do so. However, subjects in experiment 1 faced the same situation, and all of them transported the tool before any reinforcement had occurred.

Our results also differ from the phenomenon of taste aversion, whereby an animal learns within one trial to avoid the ingestion of a substance after becoming sick 1 hour later (*17*), because taste aversion involves learning to avoid a stimulus, not acquiring an instrumental response. Moreover, taste aversion shows a high degree of stimulus specificity because it develops when gustatory stimuli are paired with digestive sickness but not when gustatory stimuli are paired with other modalities of aversive stimuli. It is highly unlikely that a similar form of specialized learning also exists for tool transportation.

Another alternative is that our results represent an unlearned biological predisposition. Many species appear to plan for the future when they build nests to lay their eggs or hoard food for the winter (*18*). However, unlike tool transportation, these other activities have obvious fitness consequences (channeled by strong physiological and genetic determinants) that have favored their selection in all the individuals of a species. In contrast, tool transportation, even tool use, has not been documented in most wild populations of bonobos or orangutans (*19*, *20*). More important, apes do not store food or objects in their natural habitats because those are generally available throughout the year. It is therefore unlikely that saving and transporting tools are unlearned biological predispositions.

Because traditional learning mechanisms or certain biological predispositions appear insufficient to explain our current results, we propose that they represent a genuine case of future planning. Subjects executed a response (tool transport) that had not been reinforced during training, in the absence of the apparatus or the reward, that produced no consequences or reduced any present needs but was crucial to meet future ones. The presence of future planning in both bonobos and orangutans

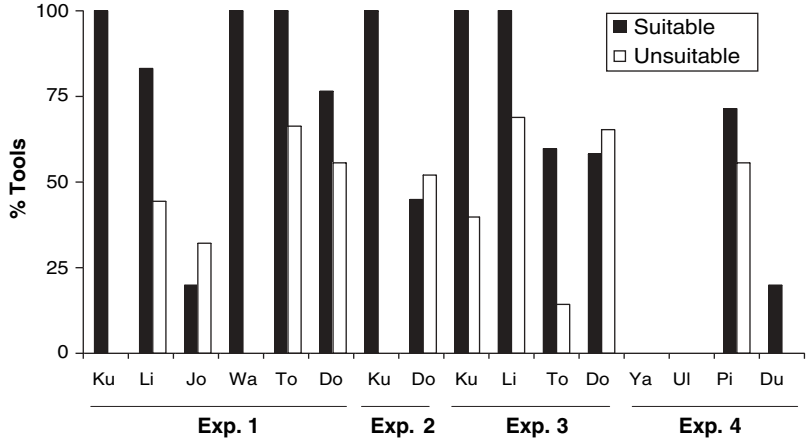


Fig. 1. Percentage of suitable and unsuitable tools transported into the test room by each subject after the delay period in each experiment. Subject name abbreviations are as follows: Ku, Kuno; Li, Limbuko; Jo, Joey; Wa, Walter; To, Toba; Do, Dokana; Ya, Yasa; Ul, Ulindi; Pi, Pini; Du, Dunja.

suggests that its precursors may have evolved before 14 Ma in the great apes. Together with recent evidence from scrub jays (*8*, *9*, *21*), our results suggest that future planning is not a uniquely human ability, thus contradicting the notion that it emerged in hominids only within the past 2.5 to 1.6 million years (*10*, *11*, *22*). Indeed, its presence in distantly related taxa such as corvids and apes reinforces the hypothesis that these taxa may have undergone convergent cognitive evolution (*23*). Future studies should investigate whether apes, like corvids, will not only transport tools for future use but also protect them from conspecifics that may steal them.

References and Notes

1. E. Tulving, in *The Missing Link in Cognition: Evolution of Self-Knowing Consciousness*, H. Terrace, J. Metcalfe, Eds. (Oxford Univ. Press, New York, 2004), pp. 3–56.
2. E. Tulving, *Behav. Brain Sci.* **7**, 223 (1984).
3. W. A. Roberts, *Psych. Bull.* **128**, 473 (2002).
4. T. Suddendorf, J. Busby, *Trends Cogn. Sci.* **7**, 391 (2003).
5. C. Boesch, H. Boesch, *Primates* **25**, 160 (1984).
6. J. Chappell, A. Kacelnik, *Anim. Cogn.* **5**, 1 (2002).
7. W. Kohler, *The Mentality of Apes* (Routledge & Kegan Paul, London, 1927).
8. N. J. Emery, N. S. Clayton, *Nature* **414**, 443 (2001).
9. N. S. Clayton, J. Bussey, A. D. Dickinson, *Nat. Rev. Neurol.* **4**, 685 (2003).
10. K. D. Schick, N. Toth, *Making Silent Stones Speak* (Simon and Schuster, New York, 1993).
11. T. Suddendorf, M. C. Corballis, *Gen. Soc. Gen. Psych. Mon.* **123**, 133 (1997).
12. G. Gergely, G. Csibra, *Inter. Stud.* **6**, 463 (2004).
13. W. J. Bailey *et al.*, *Mol. Phylogenet. Evol.* **1**, 97 (1992).
14. Detailed materials and methods are available as supporting material on *Science* Online.
15. G. A. Kimble, *Conditioning and Learning* (Appleton-Century-Crofts, New York, 1961).
16. D. A. Lieberman, D. C. McIntosh, G. V. Thomas, *J. Exp. Psych. Anim. Behav. Proc.* **5**, 224 (1979).
17. J. Garcia, D. J. Kimeldorf, R. A. Koelling, *Science* **122**, 157 (1955).
18. S. B. van der Wall, *Food Hoarding in Animals* (Univ. of Chicago Press, Chicago, IL, 1990).
19. C. P. van Schaik, E. A. Fox, L. T. Fechtman, *J. Hum. Evol.* **44**, 11 (2003).
20. T. Kano, *The Last Ape* (Stanford Univ. Press, Stanford, CA, 1992).
21. N. S. Clayton, A. D. Dickinson, *J. Comp. Psychol.* **113**, 403 (1999).
22. J. de Heinzelin *et al.*, *Science* **284**, 625 (1999).
23. N. J. Emery, N. S. Clayton, *Science* **306**, 1903 (2004).
24. We thank M. Carpenter, M. Tomasello, B. Hare, and three anonymous reviewers for their comments on a previous version of this manuscript; the keepers of the Leipzig Zoo for their help during testing; and K. Finstermeier for drawing fig. S1.

Supporting Online Material

www.sciencemag.org/cgi/content/full/312/5776/1038/DC1
Materials and Methods
Figs. S1 and S2
Table S1

26 January 2006; accepted 30 March 2006
10.1126/science.1125456

CO/FT Regulatory Module Controls Timing of Flowering and Seasonal Growth Cessation in Trees

Henrik Böhlenius,[1] Tao Huang,[1] Laurence Charbonnel-Campaa,[1] Amy M. Brunner,[2,4] Stefan Jansson,[3] Steven H. Strauss,[4] Ove Nilsson[1]*

Forest trees display a perennial growth behavior characterized by a multiple-year delay in flowering and, in temperate regions, an annual cycling between growth and dormancy. We show here that the *CO/FT* regulatory module, which controls flowering time in response to variations in daylength in annual plants, controls flowering in aspen trees. Unexpectedly, however, it also controls the short-day–induced growth cessation and bud set occurring in the fall. This regulatory mechanism can explain the ecogenetic variation in a highly adaptive trait: the critical daylength for growth cessation displayed by aspen trees sampled across a latitudinal gradient spanning northern Europe.

Trees have extended juvenile phases that can last for decades before the first flower is formed. Trees can also cycle between periods of growth and dormancy. In temperate regions, this involves a short-day–induced growth cessation and bud set in the fall, after which the tree enters a dormant state characterized by an enhanced cold tolerance. Tree populations (provenances) from northern latitudes typically display growth cessation at a longer critical daylength, leading to earlier bud set during fall compared with southern populations (*1*). This is a highly adaptive trait because it ensures that bud set and dormancy have been induced well before the risk of frost damage. This response is under strong genetic control and is maintained when trees are moved between latitudes (*2*, *3*). The molecular mechanism that controls growth cessation at different critical daylengths is not known; neither is the mechanism controlling the multiple-year delay in flowering.

In the annual plant *Arabidopsis*, the genes *CONSTANS* (*CO*) and *FLOWERING LOCUS T* (*FT*) are necessary for the daylength regulation of flowering, inducing flowering as a response to long days (*4*). *CO* displays a diurnal regulation in which the mRNA accumulation peaks at the end of the day in long days and during the night in short days (*5*). Furthermore, the CO protein is extremely labile in darkness, leading to an accumulation of CO protein only in long days (*6*). CO then induces transcription of the gene *FT* in the leaf, and the *FT* mRNA moves from leaf to shoot apex (*7*, *8*), where the translated FT protein induces the formation of flowers (*8*, *9*). The *FT* mRNA fulfills many of the criteria characterizing the elusive flower-inducing molecule "florigen" described in the 1930s (*7*).

To determine whether a tree *FT* ortholog is also involved in the regulation of flowering time in trees, a process that is not obviously regulated by daylength because of their long juvenile phase, or whether it is involved in the daylength regulation of perennial growth and dormancy, we have investigated the role of the *FT* ortholog in *Populus* trees (poplars, aspens, and cottonwoods).

We isolated the *Populus trichocarpa FT* ortholog, which we call *PtFT1* (fig. S1) and showed that its function in inducing early flowering is conserved in transgenic *Arabidopsis* (fig. S2) (*10*). *Populus trichocarpa* is difficult to transform, but all *Populus* species are closely related, and the sequence identity between homologous genes in different *Populus* species is often 99% (*11*). Male *Populus tremula x tremuloides* transformed with *35S::PtFT1* initiated flowerlike structures directly from the Agrobacterium-infected stem segments within 4 weeks (Fig. 1, A and B), compared with the normal flowering time of 8 to 20 years (*12*). This shows that *PtFT1* is a powerful inducer of flowering in *Populus*. Weaker expressing lines could be regenerated and planted in the greenhouse. These trees produced inflorescences (catkins) (Fig. 1, C to E, H, and I; and fig. S3, A to C) containing phenotypically normal male flowers (Fig. 1, F and J) with an apparently normal pollen development (Fig. 1, G and K). We also generated early-flowering female *Populus tremula* with normal inflorescence development (fig. S3D). This is the first report of juvenile transgenic trees producing inflorescences. In contrast, early-flowering *Populus* ectopically expressing

[1]Umeå Plant Science Centre, Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, S-90183 Umeå, Sweden. [2]Virginia Tech Department of Forestry, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061–0324, USA. [3]Umeå Plant Science Centre, Department of Plant Physiology, Umeå University, S-90187 Umeå, Sweden. [4]Department of Forest Science, Oregon State University, Corvallis, OR 97331–5752, USA.

*To whom correspondence should be addressed. Email: Ove.Nilsson@genfys.slu.se

the *Arabidopsis* flower meristem-identity gene *LEAFY* (*LFY*) produced shoots containing single flowers instead of catkins (*12*), reflecting the role of *LFY* as a flower meristem identity gene, while *PtFT1*, just as *FT*, appears to act as a flowering time gene, inducing more developmentally normal flowering.

The fact that *35S::PtFT1* expression leads to early flowering suggests that the level of *PtFT1* expression could be an important determinant of flowering time in *Populus* trees. To test this, we collected shoot tips from a *Populus* clone aged 2 to 6 years. This particular clone flowers first after 5 to 6 years, and floral initiation takes place between mid-May and mid-June (*13*). Samples collected in May and June displayed a gradual increase in *PtFT1* expression as the tree grew older (fig. S4), suggesting that a critical level of *PtFT1* expression is needed to initiate flowering. What kind of mechanism could be responsible for such a gradual increase over many years? In *Arabidopsis*, the gene *EARLY BOLTING IN SHORT DAYS* (*EBS*) acts as a repressor of *FT* transcription, probably through regulation of chromatin structure (*14*). It could be that the role of repressors such as *EBS* are much more pronounced in trees than in *Arabidopsis* and that each annual cycle of growth and dormancy leads to a gradual release of the chromatin structure–based *PtFT1* repression, making *PtFT1* more and more susceptible to transcriptional activation by *PtCO*.

In *Arabidopsis*, *FT*-dependent induction of flowering is the only daylength-regulated developmental transition. In contrast, shortening of the daylength induces growth cessation and bud set in *Populus* trees (*3*), and we postulated that *PtFT1* could also have a role in this daylength-regulated process. When wild-type *Populus tremula x tremuloides* plants are shifted from long days to short days, they respond by growth cessation and bud set after 32 days (Fig. 2, A and B; and fig. S5A). After another 31 days in short days and 5 days in darkness at 5°C, the tree has abscised its leaves and is dormant (Fig. 2C). In contrast, *35S::PtFT1*-expressing trees displayed no signs of growth cessation but continued growing for more than 60 days in short-day conditions (Fig. 2, E to G; and fig. S5A). This indicates that *PtFT1* expression is an efficient suppressor of short-day–induced growth cessation and bud set. We then analyzed the expression of *PtFT1* in response to a long-day to short-day shift. In long days, *PtFT1* displayed an expression pattern very similar to that of *Arabidopsis FT*, with a diurnal regulation showing a peak of expression in the beginning of the night (Fig. 2I). After transfer of plants to short days, down-regulation of *PtFT1* expression could be detected within 3 days (fig. S5B), and after one week *PtFT1* expression was severely reduced, with no signs of a diurnal variation (Fig. 2I). This indicates that the down-regulation of *PtFT1* is a very early response to a shift to short days and that this down-regulation is necessary for normal growth cessation and bud set to occur. To test this, we generated transgenic plants in which the level of *PtFT* was down-regulated through RNA interference (RNAi). These trees were more sensitive to shortening of the daylength, with nine of the *PtFT* RNAi lines setting vegetative buds 2 weeks after a shift from a 16-hour day to a 15-hour day (Fig. 2J), whereas wild-type plants required 10 weeks before buds were visible (Fig. 2J). Unlike wild-type trees, four of the RNAi lines even set buds at the 16-hour daylength (Fig. 2, D, H, and J). The increased sensitivity to the short-day signal was correlated to the level of *PtFT1* down-regulation (Fig. 2J).

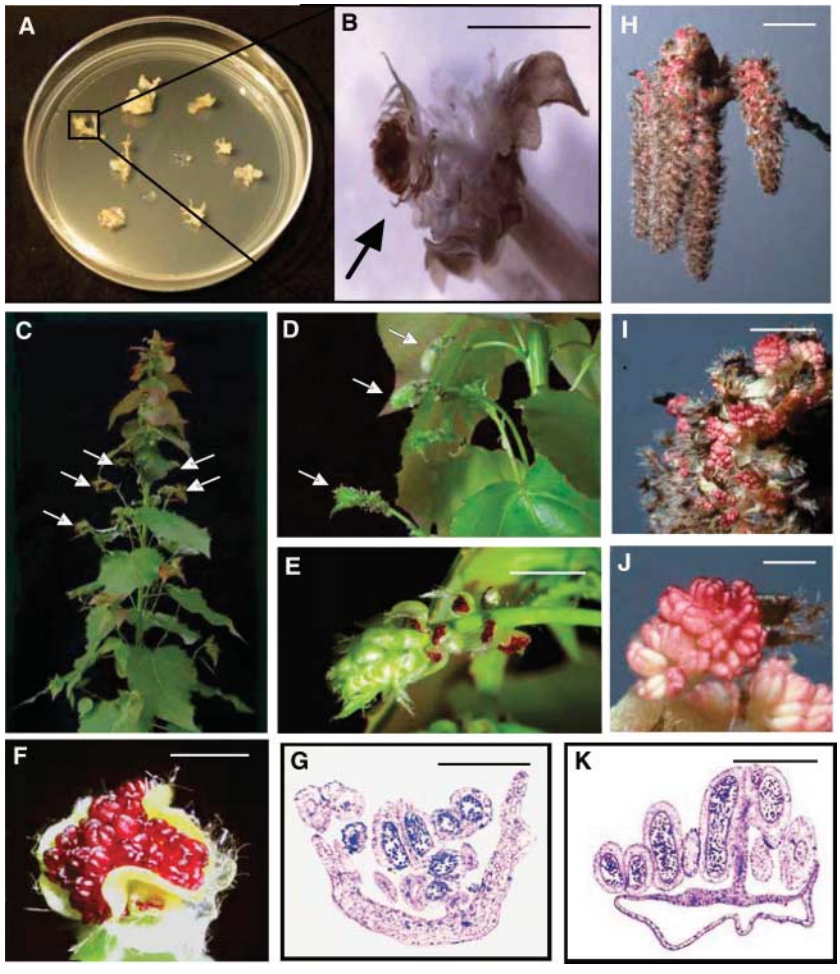


Fig. 1. *PtFT1* expression induces early flowering in *Populus* trees. (**A** and **B**) Four-week-old stem segments of hybrid aspen (*Populus tremula x tremuloides)* grown in tissue culture. (A) Overview of a tissue culture plate with stem segments transformed with *35S::PtFT1*. (B) Close-up view of a developing flower (black arrow). (**C** to **G**) Six-month-old greenhouse-grown *35S::PtFT1* plants. (C) Overview of a *35S::PtFT1* tree. (D) Developing male catkins (white arrows). (E) Close-up view of a developing male catkin. (F) Male flower removed from catkin. (G) Section of a male flower. (**H** to **K**) *Populus tremula* wild-type flowers. (H) Developing male catkins. (I) Close-up view of a developing male catkin. (J) Male flower. (K) Section of a male flower. Scale bars, 1 mm [(B), (F), (G), (J), and (K)]; 1 cm [(E), (H), and (I)].

To determine whether the regulation of *PtFT1* correlated with the critical daylengths for growth cessation displayed by European aspen (*Populus tremula*) trees originating from different latitudes (provenances), we compared trees from four different latitudes stretching from 51°N (the middle of Germany) to 63°N (Northern Sweden) (Fig. 3A). We established these plants in controlled growth conditions and determined the critical daylengths for growth cessation and bud set. They varied from 21 hours for the provenance from northern Sweden to 15 hours for the German provenance (Fig. 3B). When we grew these plants under a 19-hour photoperiod, which is above the critical daylength for the two southernmost provenances, *PtFT1* displayed a clear peak of expression at the end of the day in the southern provenances but was only very weakly expressed in the two

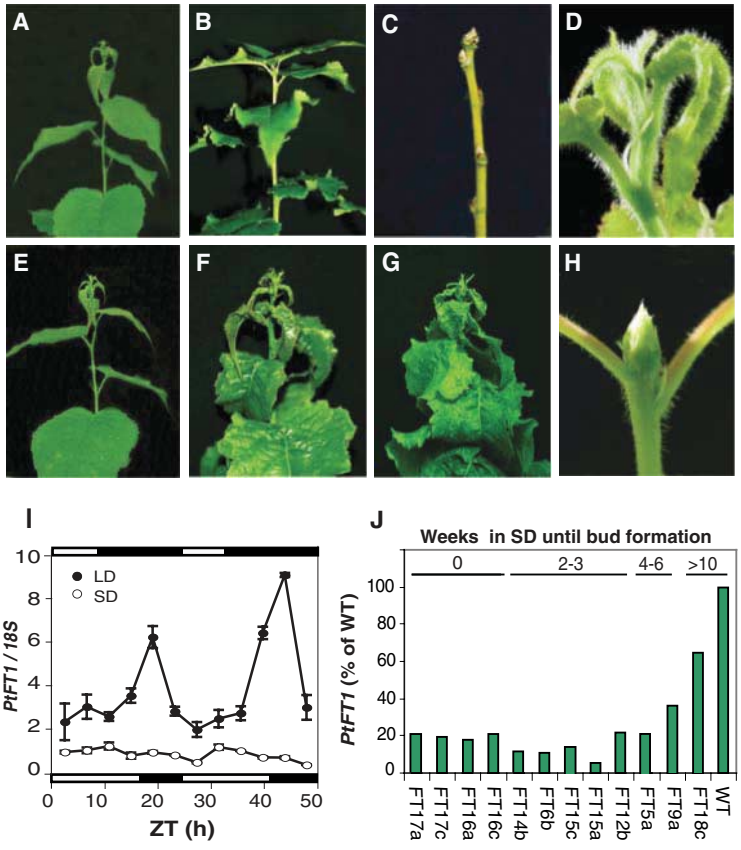


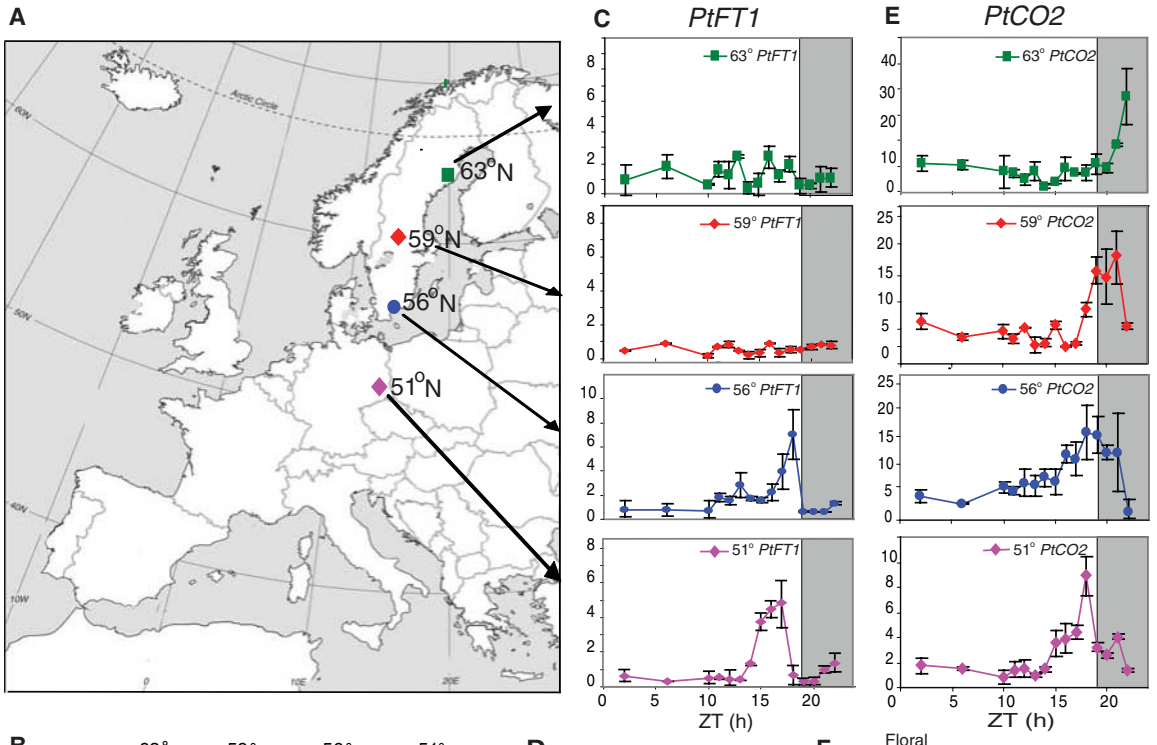
Fig. 2. *PtFT1* expression affects short-day–induced growth cessation and bud set in *Populus* trees. (**A** to **C**) Wild-type plants and (**E** to **G**) *35S::PtFT1 Populus tremula x tremuloides* in long days [(A) and (E)], 32 short days [(B) and (F)], and 63 short days and 5 days in darkness at 5°C [(C) and (G)]. (**D**) Wild-type and (**H**) *PtFT* (*PtFT1* and *PtFT2*) RNAi plants grown in 16-hour-long days. (**I**) Diurnal expression pattern of *PtFT1* in long days (closed circles) and 1 week after transfer to short days (open circles). Three mature leaves were collected at the different time points from three independent *Populus tremula x tremuloides* (T89) ramets. Black boxes indicate darkness and white boxes light. Error bars, ± SD. ZT, zeitgeber time. (**J**) *PtFT1* expression in 12 independent transgenic *PtFT* (*PtFT1* and *PtFT2*) RNAi lines setting bud 0 to 10 weeks after a shift from 16-hour to 15-hour days (5 ramets per clone). All ramets of the same clone set bud within the indicated time intervals. Three mature leaves were collected and pooled in the evening from three ramets of each clone growing in 16-hour days. Details of these experiments are described in (*10*).

northern provenances (Fig. 3C). Further experiments at 17- and 21-hour daylength confirmed that *PtFT1* only showed a peak in expression when the photoperiod was above the critical daylength for a particular provenance (fig. S6A), supporting the conclusion that the level of *PtFT1* expression is a critical determinant of the timing of growth cessation and bud set in the fall.

What, then, is the mechanism sensing these different critical daylengths? In *Arabidopsis*, *FT* transcription is under the control of the gene *CO* which, according to the external coincidence model, is responsible for *FT* regulation in response to long days (*15*, *16*). We therefore tested whether the function of the *CO*/*FT* regulon was conserved in *Populus* and whether variations in expression of the *Populus CO* ortholog could explain the differential regulation of *PtFT1* seen in the different provenances. We isolated a *Populus trichocarpa CO* ortholog, *PtCO2* (fig. S1) (*10*). In long days, *PtCO2* displays a diurnal expression pattern peaking at the end of the day. After a shift to short days, the expression peaks in the night, with very low expression in the light (Fig. 3D). This is very similar to the expression pattern of *CO* and corresponds to the *Arabidopsis* model of *FT* regulation by *CO*, in which *CO* can only activate *FT* transcription when *CO* transcription peaks at the end of the day (*15*, *16*). An attractive model explaining the different critical daylengths for growth cessation and bud set for the various provenances would be that the *PtCO2* expression starts increasing earlier after dawn in the more southern provenances. This would mean that the days must be shorter for the southern than for the northern provenances before *PtCO2* transcript abundance starts increasing in darkness and no *PtFT1* expression is induced. In a 19-hour photoperiod, this is indeed the case (Fig. 3E). In the southernmost provenance, the first detectable rise in *PtCO2* transcript accumulation occurs about 6 hours earlier than in the northernmost provenance (Fig. 3E); this is also the difference in critical daylength between these provenances (Fig. 3B). These changes in the phase of *PtCO2* mRNA oscillation are also seen at 17-hour and 21-hour daylengths (fig. S6B). The importance of the level of *PtCO* transcription for the timing of growth cessation is further corroborated by the fact that transgenic plants with a down-regulation of *PtCO* expression are more sensitive to the short-day signal, just like *PtFT* RNAi plants (fig. S7).

In *Populus*, phytochrome genes have been implicated in the regulation of short-day–induced bud set and growth cessation (*2*, *17*, *18*), and transgenic *Populus* trees expressing the oat *PHYTOCHROME A* (*PHYA*) gene fail to initiate growth cessation and bud set as a response to the short-day signal (*19*). Because PHYA is a known regulator of *FT* transcription through the regulation of *CO* (*16*), we tested whether one reason that *35S::PHYA* plants fail to respond to the short-day signal could be an inability to down-regulate *PtFT1*. In our conditions, the *35S::PHYA* plants continued growing after being transferred to short days, much in the same way as did *35S::PtFT1* plants (fig. S5A). Furthermore, the *35S::PHYA* plants displayed no signs of down-regulating *PtFT1* during short days (fig. S5B), and the amount of *PtCO2* mRNA did not decrease at the end of the day, suggesting a shift in the phase of *PtCO2* expression (fig. S5C). Taken together, these results suggest that the level of *PtFT1* transcription is an important regulator of short-day-induced growth cessation and bud set and that the phytochrome regulation of critical daylength could, at least partly, be mediated through a regulation of *PtCO2* and *PtFT1*.

We show in these experiments that *PtFT1*, the *Populus* ortholog of the *Arabidopsis* gene *FT*, is controlling two aspects of the perennial growth behavior. It is involved both in controlling the multiple-year delay in flowering time and in controlling growth cessation and bud set in the fall. In a reproductively mature tree, both these processes might be subjected to daylength control—flowering induced by long days in the spring and early summer, and growth cessation induced in the fall. Indeed, this is reflected in the *PtFT1* expression pattern over a season (Fig. 3F). Our results therefore suggest that *FT* orthologs can have a more



B

	63°	59°	56°	51°
22 h	+	+	+	+
21 h	-	+	+	+
20 h	-	+	+	+
19 h	-	-	+	+
18 h	-	-	+	+
17 h	-	-	-	+
15 h	-	-	-	-

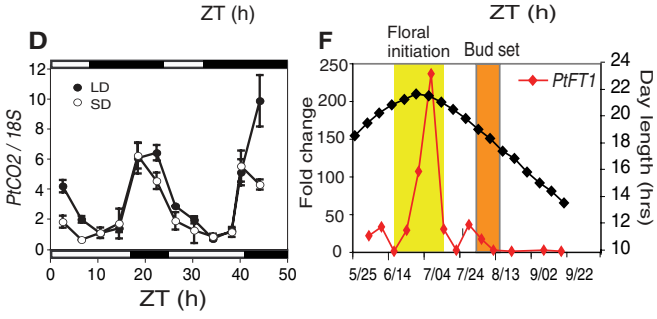


Fig. 3. The critical daylength for growth cessation in trees originating from different latitudes is controlled by *PtFT1* expression and the phase of *PtCO2* oscillation. (**A**) Map of Europe indicating the origin of the four different European aspen (*Populus tremula*) clones. Green squares, Umeå (63°N); red diamonds, Brunsberg (59°N); blue circles, Ronneby (56°N); lilac diamonds, Brauna (51°N). (**B**) Table showing the critical daylengths for growth cessation for the collected *Populus tremula* clones. Plus indicates growth, minus indicates bud set. (**C**) *PtFT1* and (**E**) *PtCO2* expression in the different *Populus tremula* clones in 19-hour daylength. Gray boxes indicate night, white boxes indicate light. (**D**) Diurnal expression pattern of *PtCO2* in *Populus tremula x tremuloides* (T89) under long days (filled circles) and 1 week after transfer to short days (open circles). Three mature leaves were collected at the different time points from three independent ramets of each clone. Black boxes indicate darkness and white boxes light. Error bars, ±SD. ZT, zeitgeber time. (**F**) Seasonal variation in *PtFT1* expression (red diamonds) in a 30-year-old *Populus tremula* tree growing in Umeå, Sweden (63°N). Black diamonds indicate daylength. Shaded areas indicate bud set and the likely period for floral initiation. Three mature leaves were collected and pooled in the middle of the day at the dates indicated. Details of these experiments are described in (*10*).

general role in regulating biological processes that are subjected to daylength control than previously anticipated from *Arabidopsis* work.

We also show a mechanism for how the *CO*/*FT* regulon is controlling a highly adaptive trait for forest trees: the variation in the critical daylength that induces growth cessation and bud set in tree populations originating from different latitudes. This response represents a critical ecological and evolutionary tradeoff between survival and growth in most forest trees (*20*–*22*) and is key to the adaptation to new geographical areas and to the current climate change. Knowledge about this mechanism will therefore be of importance for future tree breeding programs. The technique to induce early flowering also opens up the possibility to dramatically speed up tree breeding programs, which is greatly needed because the world's natural forests are declining partly as a result of an increasing demand for wood.

References and Notes

1. S. S. Pauley, T. O. Perry, *J. Arnold Arbor. Harv. Univ.* **35**, 167 (1954).
2. B. E. Frewen *et al.*, *Genetics* **154**, 837 (2000).
3. G. T. Howe, G. Gardner, W. P. Hackett, G. R. Furnier, *Physiol. Plant.* **97**, 95 (1996).
4. M. Koornneef, C. J. Hanhart, J. H. van der Veen, *Mol. Gen. Genet.* **229**, 57 (1991).
5. P. Suarez-Lopez *et al.*, *Nature* **410**, 1116 (2001).
6. F. Valverde *et al.*, *Science* **303**, 1003 (2004).
7. T. Huang, H. Böhlenius, S. Eriksson, F. Parcy, O. Nilsson, *Science* **309**, 1694 (2005).
8. P. A. Wigge *et al.*, *Science* **309**, 1056 (2005).
9. M. Abe *et al.*, *Science* **309**, 1052 (2005).
10. Materials and methods are available as supporting material on *Science* Online.
11. P. K. Ingvarsson, *Genetics* **169**, 945 (2005).
12. D. Weigel, O. Nilsson, *Nature* **377**, 495 (1995).
13. W. H. Rottmann *et al.*, *Plant J.* **22**, 235 (2000).
14. M. Pineiro, C. Gomez-Mena, R. Schaffer, J. M. Martinez-Zapater, G. Coupland, *Plant Cell* **15**, 1552 (2003).
15. I. Searle, G. Coupland, *EMBO J.* **23**, 1217 (2004).
16. M. J. Yanovsky, S. A. Kay, *Nature* **419**, 308 (2002).
17. G. T. Howe *et al.*, *Mol. Biol. Evol.* **15**, 160 (1998).
18. P. K. Ingvarsson, V. Garcia, D. Hall, V. Luquez, S. Jansson, *Genetics* **172**, 1845 (2006).
19. J. E. Olsen *et al.*, *Plant J.* **12**, 1339 (1997).
20. C. J. Weiser, *Science* **169**, 1299 (1970).
21. D. P. Horvath, J. V. Anderson, W. S. Chao, M. E. Foley, *Trends Plant Sci.* **8**, 534 (2003).
22. G. T. Howe *et al.*, *Can. J. Bot.* **81**, 1247 (2003).

23. We thank H. Hönicka for the gift of the *Populus tremula* "Brauna" clone, T. Moritz for the *35S::PHYA* plants, O. Shevchenko for technical assistance, and V. Hurry for critical reading of the manuscript. This work was supported by a grant through Individual Grants for the Advancement of Research Leaders from the Swedish Foundation for Strategic Research (O.N.), grants from the Swedish Research Council (S.J.), and USDA Cooperative State Research, Education, and Extension Service grant 2002-35301-12173 (A.M.B.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/1126038/DC1
Materials and Methods
SOM Text
Figs. S1 to S7
Table S1
References

9 February 2006; accepted 24 April 2006
Published online 4 May 2006;
10.1126/science.1126038
Include this information when citing this paper.

Emergent Properties of Reduced-Genome *Escherichia coli*

György Pósfai,[1,2*] Guy Plunkett III,[2,3,4] Tamás Fehér,[1] David Frisch,[2,4] Günther M. Keil,[5] Kinga Umenhoffer,[1] Vitaliy Kolisnychenko,[1†] Buffy Stahl,[2] Shamik S. Sharma,[6‡] Monika de Arruda,[2] Valerie Burland,[2,3] Sarah W. Harcum,[7] Frederick R. Blattner[2,3,4*]

With the use of synthetic biology, we reduced the *Escherichia coli* K-12 genome by making planned, precise deletions. The multiple-deletion series (MDS) strains, with genome reductions up to 15%, were designed by identifying nonessential genes and sequences for elimination, including recombinogenic or mobile DNA and cryptic virulence genes, while preserving good growth profiles and protein production. Genome reduction also led to unanticipated beneficial properties: high electroporation efficiency and accurate propagation of recombinant genes and plasmids that were unstable in other strains. Eradication of stress-induced transposition evidently stabilized the MDS genomes and provided some of the new properties.

Escherichia coli K-12 is one of the best understood and most thoroughly analyzed organisms and is the platform of choice for genetic, biochemical, and metabolic simulation research. Commercially, it is used for production of metabolites such as amino acids and proteins of therapeutic or commercial interest. K-12 is also gaining ground for production of DNA for gene therapy, DNA vaccines, and interference RNA. The genomes of two closely related K-12 strains, MG1655 and W3110, have been sequenced (*1*–*3*), and 87% of their genes have functional assignments (*4*). Because *E. coli* evolved in animal intestines and in the environment, parts of its genome are unnecessary for some applications, possibly even counterproductive. By eliminating as many of these gene segments as possible, we have constructed genetically stable "tabula rasa" strains with robust metabolic performance, to which genes for practical applications may be added.

Genome reductions may improve metabolic efficiency and decrease the redundancy among *E. coli* genes and regulatory circuits. Disseminated throughout the genome are mobile DNA elements that mediate recombination events such as transposition and horizontal gene transfer, including insertion sequence (IS) elements, transposases, defective phages, integrases, and site-specific recombinases (*5*). Multiple elements also provide DNA sequence repeats that mediate inversions, duplications, and deletions by homologous recombination even without transposase. To stabilize the genome and streamline metabolism, these elements must be deleted and unwanted functions removed, such as those specific for human hosts or particular environments. By means of a rational design strategy, we avoided loss of robustness that would result from more extensive deletions or an attempt to construct a minimal genome.

Predicting genes to be deleted without detrimental effect is not trivial. MG1655 reduced by 29.7% (*6*) had severely impaired growth and chromosomal segregation, whereas a strain reduced by 7% grew normally (*7*). We used a series of genomic sequence comparisons (Fig. 1) to identify segments present in K-12 but absent from five other *E. coli* (*8*). The analysis yielded nearly 100 proposed deletions (20% of the genome), encoding 900 genes. Initially we targeted large islands, IS-containing islands, and individual genes containing IS elements for removal. Deletion methods were based on recombination mediated by the phage lambda Red system. Beginning with prototype strain MDS12 (*9*), "scarless" deletions were made by removing the targeted segment and resealing the genome so that markers used in the construction were eliminated. Resulting strains were tested for robust growth on minimal medium, and deletions were serially accumulated into a single strain by P1 transduction. Deletion endpoints were verified by sequencing and by DNA microarray hybridization (Fig. 1) (*8*). Physical characteristics of the MDS strains are summarized in Table 1; deletion endpoints are in table S1, deleted genes in table S3, and strain request information in (*8*). Generation of double-strand breaks (DSBs) in each deletion process might have induced error-prone repair, but experiments designed to detect this showed that a single transient break would have no detectable effect on the accumulation of point mutations.

MDS39, the first in the series designed to be IS-free, was examined by genomic DNA hybridization to NimbleGen genome scanning microarrays, which included IS elements, phages, and plasmids absent from K-12 (*8*) as well as the K-12 genomic sequence in the form of 24-base oligonucleotides tiled about every 50 bases on both strands. Alarmingly, we found five unexpected copies of IS that had transposed to new locations (*8*) since the project began in 2002. Specific deletions later removed these IS and other segments, resulting in MDS41, 42, and 43 (*8*).

The reduced strains functioned comparably to the parent, MG1655. Growth rates were very similar (*8*). The slight changes in replichore lengths (Table 1) had no impact. Electroporation efficiencies of DH10B, MG1655, and MDS42 were compared (Table 2) for a small multicopy plasmid (pUC19) and pCC145, a bacterial artificial chromosome with a 145-kb insert. The efficiency for MDS42 was two orders of magnitude higher than that of MG1655 ($P = 0.002$), comparable to that of DH10B (regarded as best for electroporation). MDS42 efficiencies equaled or exceeded those of purchased competent cells, both under conditions optimal for MDS and according to the manufacturer's protocol

Table 1. Summary of the deletion strains and MG1655, based on the updated MG1655 sequence (U00096.2, June 2004); MDS12 values include the original MD1 deletion (see table S1). Gene counts are based on recently updated annotations (*3*).

	MG1655	MDS12	MDS41	MDS42	MDS43
Total number of genes	4,434	4,011	3,731	3,730	3,691
Genome size (bp)	4,639,675	4,263,492	3,977,067	3,976,359	3,931,408
Replichore imbalance (bp)	30,517	141,360	139,331	138,623	183,574
Total no. genes deleted	0	423	703	704	743
Total bp DNA deleted	0	376,183	662,608	663,316	708,267
Percent of genome deleted	0	8.11%	14.28%	14.30%	15.27%

[1]Institute of Biochemistry, Biological Research Center, H-6726 Szeged, Hungary. [2]Scarab Genomics LLC, Madison, WI 53713, USA. [3]Department of Genetics, University of Wisconsin, Madison, WI 53706, USA. [4]Genome Center of Wisconsin, University of Wisconsin, Madison, WI 53706, USA. [5]Federal Research Centre for Virus Diseases of Animals, Institute of Molecular Biology, Friedrich-Loeffler Institutes, D-17493 Greifswald-Insel Riems, Germany. [6]Department of Chemical and Biomolecular Engineering, Clemson University, Clemson, SC 29634, USA. [7]Department of Bioengineering, Clemson University, Clemson, SC 29634, USA.

*To whom correspondence should be addressed. E-mail: fred@genome.wisc.edu (F.R.B.); posfaigy@nucleus.szbk.u-szeged.hu (G.P.)

†Present address: National Heart, Lung, and Blood Institute, 9000 Rockville Pike, Building 10, Room 7D05, Bethesda, MD 20852, USA.

‡Present address: New England Biolabs, 240 County Road, Ipswich, MA 01938, USA.

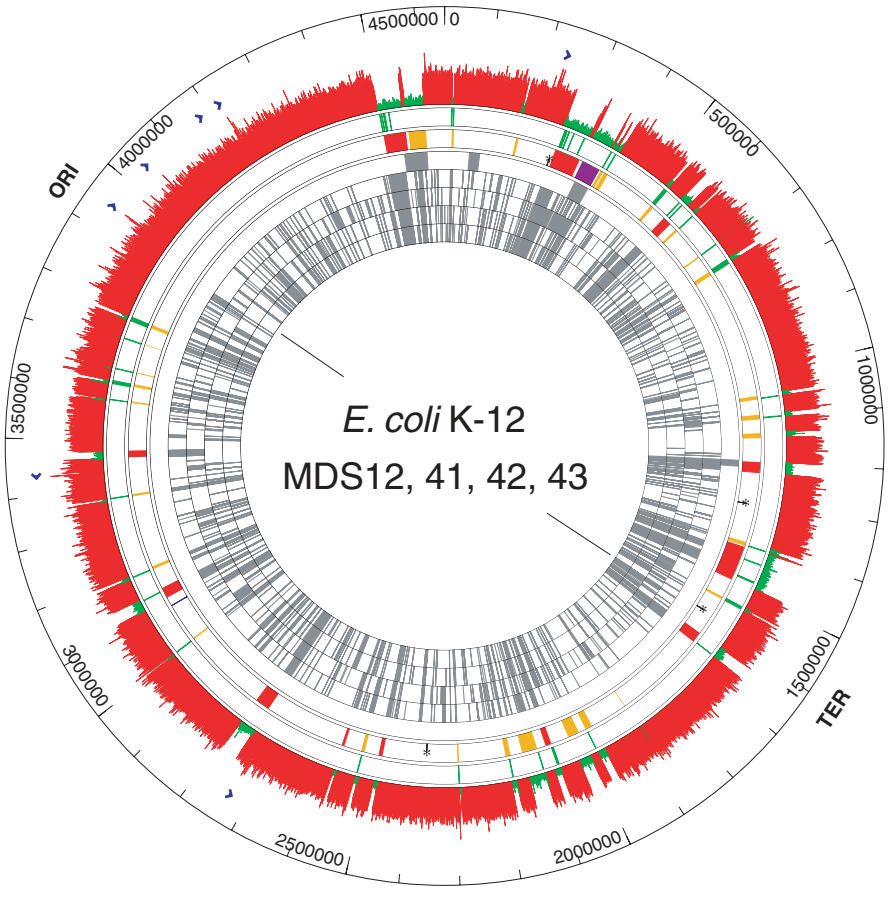


Fig. 1. Design and validation of MDS. Rings depict features mapped to the genome of *E. coli* K-12 strain MG1655, numbered on the outer ring. Outward from the center (rings 1 to 5; gray) are regions of K-12 that are absent in other *E. coli* genomes: RS218, CFT073, *S. flexneri* 2457T, O157:H7 EDL933, and DH10B. Ring 6: Regions targeted for deletion (red, MDS12; yellow, MDS41; blue, MDS42; purple, MDS43; asterisks, IS elements detected in MDS39 and later removed). Ring 7: Native IS and Rhs repeat elements (green). Ring 8: Experimental confirmation of the deletions in MDS43 by a genome-scanning DNA microarray (green, probes corresponding to deletions; red, other probes). Outer ring: ORI and TER, origin and terminus of replication, respectively; rRNA operons are in blue. See (*8*) for further details.

for DH10B. Chemically competent (*10*) MDS42 performed similarly to DH10B (*8*). In fermentations, MDS strains grew to high cell densities by a fed-batch protocol on minimal medium (Fig. 2A). Recombinant protein expression for the model protein chloramphenicol acetyltransferase (CAT) was similar for MDS41 and MG1655 grown to high cell densities (Fig. 2B). An exogenous DNA methyltransferase was also expressed efficiently in MDS42 but displayed low yields in an undeleted host (*11*).

To track IS transposition during experimental procedures, we examined plasmid DNAs isolated from different hosts (*8*). When purified from hosts with IS elements in their genomes, the plasmids were frequently contaminated with elements carried in (or that cotransform with) the plasmid DNA (fig. S1). Elements present in the host were detectable in transformants, even when purchased plasmids were used.

To verify that MDS strains are free from IS-mediated mutagenesis, we examined mutant bacteria that gain spontaneously the ability to use salicin as the sole carbon source. Metabolism of salicin by *E. coli* requires activation of the *bgl* operon, which occurs primarily by IS insertion into the promoter region (*12*). When MDS41 and MG1655 mutants were selected with salicin as the sole carbon source, the activation rate for MDS41 was less than 8% of that for MG1655 (Fig. 3A). The polymerase chain reaction (PCR) confirmed the absence of IS-generated mutations in MDS41, whereas numbers of IS-unrelated mutations were the same in both strains (fig. S4A).

We used classical fluctuation assays to analyze point mutations, deletions, and insertions in growing populations of bacteria by selection of mutants resistant to D-cycloserine, an antibiotic imported by CycA permease (*13*, *14*). Resistance arises almost

Table 2. Electroporation efficiency of MDS42 and undeleted *E. coli* strains with large single-copy and small multicopy plasmids (*8*).

Strain	Plasmid	Transformants/μg
MG1655	pUC19†	0.7×10^8
DH10B	pUC19	35.0×10^8
DH10B*	pUC19	35.4×10^8
MDS42	pUC19	130.0×10^8
MG1655	pCC145‡	0
DH10B	pCC145	1.9×10^6
DH10B*	pCC145	6.5×10^6
MDS42	pCC145	10.0×10^6

*Purchased competent cells. †2.686 kb. ‡153 kb.

exclusively from loss-of-function mutations in *cycA*, which do not affect growth in minimal medium (*15*). The total mutation rates of *cycA* in MG1655 and MDS41 were 6.56×10^{-8} and 5.27×10^{-8}, respectively; this difference (21.2%) was significant ($P \leq 0.0001$; table S2) (*8*). In MG1655, PCR revealed that IS transpositions accounted for 24.2% of the mutations, deletions accounted for 1.5%, and point mutations and small indels accounted for 74.3% (Fig. 3B). In MDS41, no IS-related mutations were found, and frequencies of other mutation types were similar to those in MG1655; thus, the differences are explained by the absence of IS elements in MDS41.

Recombinant ectopic genes are not always tolerated by *E. coli*, and IS mutagenesis provides a defense against expression of products that are deleterious. A chimeric gene composed of VP60 of rabbit hemorrhagic disease virus (*16*) fused to the B subunit of cholera toxin (CTX) was very unstable in *E. coli*. Individually, both genes carried by a low copy number plasmid were stable in *E. coli* HB101, C600, and DH10B, but pCTXVP60 carrying the fusion gene in the same hosts did not produce fusion protein and was recovered in low yields. All recovered plasmids contained mutations in the CTXVP60 open reading frame, virtually all resulting from IS insertions (fig. S2). In contrast, the recombinant plasmid was completely stable in MDS; normal yields of plasmid DNA were obtained.

Plasmids based on adeno-associated virus are used as delivery vehicles in vaccine and gene therapy research (*17*). They are unstable when propagated in standard *E. coli* hosts. The plasmid pT-ITR contains both inverted terminal repeat sequences (ITRs) of the virus. The ITRs fold in perfectly paired, stable secondary structures with double arms (fig. S3A) that frequently delete in *E. coli* (*18*), necessitating extensive screening for intact plasmids before use in gene therapy. We tested pT-ITR stability in MG1655 and MDS42 over serial subcultures (fig. S3B). When grown in MG1655, plasmid restriction

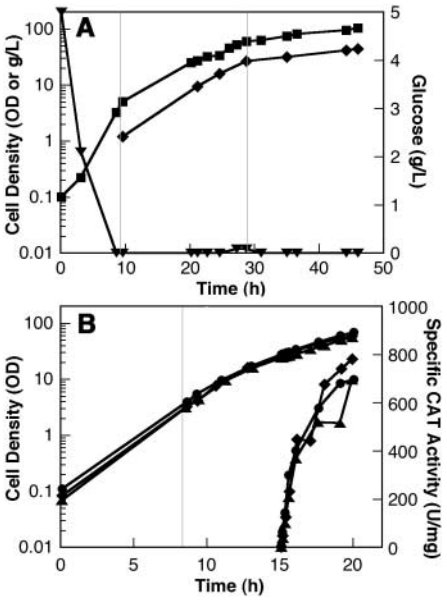


Fig. 2. Growth and protein production. (**A**) MDS41 in minimal medium. Three growth phases (phase 1, batch phase; phase 2, fed-batch, controlled growth rate 0.15 hour^{-1}; phase 3, fed-batch, controlled growth rate 0.03 hour^{-1} to avoid oxygen limitation), marked by vertical lines, were used to reach a dry cell weight (DCW) of 44 g/liter, optical density at 600 nm (OD_{600}) > 100. ■, optical density (left scale); ♦, DCW (left scale); ▼, glucose concentration (right scale). (**B**) Cell density and CAT expression from pProEX HT-CAT in MG1655 and MDS41 in minimal medium, a single fed-batch phase controlled to 0.25 hour^{-1}. Isopropyl-β-D-thiogalactopyranoside (IPTG; 5 mM) was added at 15 hours to induce CAT expression. ●, MG1655; ■ and ▲, MDS41 duplicates.

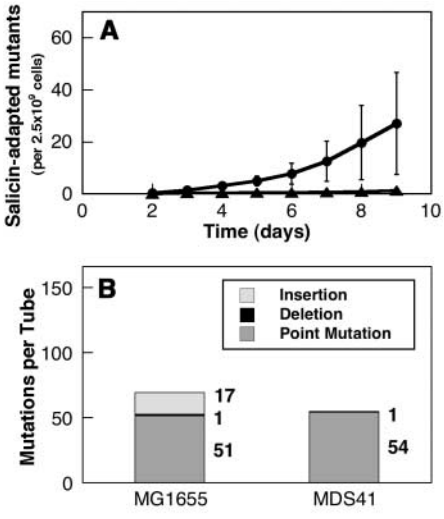


Fig. 3. Mutation rates and spectrum. (**A**) Adaptation of MG1655 (●) and MDS41 (▲) cells to salicin/minimal medium. (**B**) *cycA* mutations causing D-cycloserine resistance in MG1655 and MDS41. Total cell numbers and SD values are in table S2 (*8*).

digests produced multiple new DNA fragments, whereas propagation in MDS42 produced uniform digest patterns over several subcultures. The digest patterns indicated loss of both ITRs in MG1655, which we confirmed by DNA sequencing (*8*).

We showed that pCTXVP60 is not a specific target for IS transposition by performing bystander mutation assays (*8*) on MG1655 with pCTX or pCTXVP60, for mutations to salicin adaptation and D-cycloserine resistance (CS^R). Bacteria containing pCTXVP60 showed a rate of *bgl* mutations 4 times that seen in bacteria containing pCTX (fig S4B). Fluctuation assays also showed higher rates of CS^R when pCTXVP60 was present (fig. S4C, table S2B), and one-third of *cycA* insertion mutants also had IS in the plasmid. No transposition was detected in the plasmid-encoded CAT gene when expression was induced (table S2C), but CS^R insertion mutants were more than twice as frequent as they were without induction (fig. S5).

In this work, deletions totaling up to 15.27% of the *E. coli* genome produced stable strains without physiological compromise. Elimination of transposition artifacts was expected, but neither the increased electroporation efficiency nor the stability of plasmids unstable in K-12 was anticipated. The unexpectedly efficient electrocompetence of MDS42 counters the suggestion (*19*) that a *deoR* mutation in DH10B is the critical determinant for high electroporation efficiency; both MDS and DH10B (*20*) are *deoR*$^+$.

Removal of external structures such as fimbriae could allow better access for DNA to the depolarized membrane, but removal of an unknown deoxyribonuclease or restriction system or activation of an unknown DNA uptake factor could also affect the recovery of transformants. More than 180 of the genes deleted from MDS encode known or predicted membrane-associated proteins (e.g., fimbrial and flagellar structures, transport systems), membrane synthesis enzymes, or regulatory factors, all of which could influence membrane composition cumulatively to bring about altered sensitivity to depolarization. Because much of the K-12 protein interactome remains obscure, unexpected results of multiple deletions are likely. The synthetic biology of genome reduction could have produced synergistic interactions ("synthetic beneficials") between deletions that resulted in an altered phenotype such as high electrocompetence, whereas other combinations of deletions would result in less surprising synthetic lethals.

All strains tested, except MDS, were affected by contamination of isolated plasmids with IS-containing DNA from carryover of IS-containing genomic DNA, IS mini-circles (*21*), and plasmids carrying integrated IS. Our results show that any DNA propagated on *E. coli* containing IS elements is likely to be contaminated with IS elements, and that transposition can be frequent. IS transposition mutates both plasmid and chromosomal bystander genes in a manner consistent with stress-induced activity; powerful selection must also operate. Removing the IS eliminated the main source of instability.

References and Notes

1. F. R. Blattner *et al.*, *Science* **277**, 1453 (1997).
2. K. Hayashi *et al.*, *Mol. Syst. Biol.*, published online 21 February 2006 (10.1038/msb4100049).
3. M. Riley *et al.*, *Nucleic Acids Res.* **34**, 1 (2006).
4. M. H. Serres, S. Goswami, M. Riley, *Nucleic Acids Res.* **32**, 300D (2004).
5. L. S. Frost, R. Leplae, A. O. Summers, A. Toussaint, *Nat. Rev. Microbiol.* **3**, 722 (2005).
6. M. Hashimoto *et al.*, *Mol. Microbiol.* **55**, 137 (2005).
7. B. J. Yu *et al.*, *Nat. Biotechnol.* **20**, 1018 (2002).
8. See supporting material on *Science* Online.
9. V. Kolisnychenko *et al.*, *Genome Res.* **12**, 640 (2002).
10. H. Inoue, H. Nojima, H. Okayama, *Gene* **96**, 23 (1990).
11. A. Kiss, personal communication.
12. B. G. Hall, *Mol. Biol. Evol.* **15**, 1 (1998).
13. R. J. Wargel, C. A. Hadur, F. C. Neuhaus, *J. Bacteriol.* **105**, 1028 (1971).
14. R. R. Russell, *J. Bacteriol.* **111**, 622 (1972).
15. T. Fehér, B. Cseh, K. Umenhoffer, I. Karcagi, G. Pósfai, *Mutat. Res.* **595**, 184 (2006).
16. G. Meyers, C. Wirblich, H. J. Thiel, *Virology* **184**, 664 (1991).
17. L. Santat *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 11053 (2005).
18. S. Chatterjee, personal communication.
19. S. G. Grant, J. Jessee, F. R. Bloom, D. Hanahan, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4645 (1990).
20. Our analysis of preliminary genome sequence data for DH10B obtained from the Human Genome Sequencing Center (www.hgsc.bcm.tmc.edu/projects/microbial/EcoliDH10B).
21. Y. Sekine, N. Eisaki, K. Kobayashi, E. Ohtsubo, *Gene* **191**, 183 (1997).
22. We thank K. Giesow, A. Szalkanovics, N. Zimmerman, N. Hermersmann, S. Phillips, E. Steffen, and J. Altmann for excellent technical contributions; M. Chandler for sequences of IS elements; I. Karcagi for help with DNA constructs; J. Luecke for help with microarrays; and T. Albert, S. C. Kim, and J. Campbell for critical reading of the manuscript. Supported by a grant from the Hungarian Scientific Research Fund (OTKA) (G.P.), NIH grant GM35682 (F.R.B.), and the Federal Ministry of Education and Research (G.M.K.). F.R.B. is president and CEO of and has a financial interest in Scarab Genomics, LLC, is president of and has a financial interest in DNASTAR, Inc., and is on the scientific advisory board of and has a financial interest in Nimblegen Systems, Inc.

Supporting Online Material

www.sciencemag.org.cgi/content/full/1126439/DC1
Materials and Methods
SOM Text
Figs. S1 to S5
Tables S1 to S4
References and Notes

21 February 2006; accepted 19 April 2006
Published online 27 April 2006;
10.1126/science.1126439
Include this information when citing this paper.

Agent-Specific Responses in the Cingulate Cortex During Economic Exchanges

Damon Tomlin,[1*] M. Amin Kayali,[1*] Brooks King-Casas,[1] Cedric Anen,[3] Colin F. Camerer,[3] Steven R. Quartz,[3] P. Read Montague[1,2]

Interactions with other responsive agents lie at the core of all social exchange. During a social exchange with a partner, one fundamental variable that must be computed correctly is who gets credit for a shared outcome; this assignment is crucial for deciding on an optimal level of cooperation that avoids simple exploitation. We carried out an iterated, two-person economic exchange and made simultaneous hemodynamic measurements from each player's brain. These joint measurements revealed agent-specific responses in the social domain ("me" and "not me") arranged in a systematic spatial pattern along the cingulate cortex. This systematic response pattern did not depend on metrical aspects of the exchange, and it disappeared completely in the absence of a responding partner.

Social exchange occurs in species ranging from insects to humans (*1–3*). In primates, reciprocal interactions with nonkin occur repeatedly, thus necessitating the capacity to assign social credit or blame for shared outcomes and to act appropriately according to these assignments (*4–6*). In humans, reciprocity is a central feature of the collection of psychological mechanisms necessary to support social exchange (*3*); yet, the underlying neural representations of these mechanisms remain murky. In almost all social exchanges, one must detect and accurately track which social agent (who) gets credit for an outcome. Should credit for an outcome be assigned to one's own actions or those of one's partner? Perhaps such assignments are more a matter of degree—assigning the degree-of-credit to some shared outcome. Understanding such agent-specific mechanisms is important, because the assignment of social agency (*7–13*) appears to break down in a range of mental illnesses (*14–16*).

Social agency computations are also a prerequisite for generating models of others's mental states. This latter capacity, called theory-of-mind, is highly developed in humans and has been shown to activate a consistent set of brain regions in neuroimaging experiments (*17–20*). Recent work has complemented these theory-of-mind experiments by using interactive economic games as ecologically realistic models for human exchange (*21–31*). These experiments have elicited not only brain responses in previously described theory-of-mind networks (*27–29*), but also have elicited formerly unreported activations along the cingulate cortex that correlate with the revelation of a social partner's decision (*29*). Although evoked during an economic exchange with another human, these cingulate activations did not modulate as a function of the fairness of the exchange, nor did they occur in exchanges with computer partners (*28*).

This lack of sensitivity to measures of outcome suggests that these responses do not encode some metrical aspect of the trade; instead, they are consistent with the social agency computation described above. To test this possibility directly, we scanned both brains of two subjects interacting in a 10-round trust game (*32*, *33*). This game has been used to identify neural correlates of reputation building and reciprocity (*25*). Because the trust game is a multiround economic exchange, it allows us to estimate brain responses in both subjects to multiple revelations of a partner's choice and multiple instances of one's own choice.

The trust game is illustrated in Fig. 1A. In each round, one player (investor) invests some amount I (investment phase) that is tripled and sent to the other player (trustee) who decides to repay some fraction f of the tripled amount (repayment phase) (*33*). Players maintain their roles for 10 consecutive rounds, and information about outcomes within each phase is presented simultaneously to both subjects. We parameterized social context within the trust game by implementing two separate versions of the task in different subject cohorts: personal (n = 104 subjects) and impersonal [n = 96; previously reported in (*25*)]. In the personal version, subjects met before the task, were instructed together, saw a picture of their partner during each round of the game, and met their partner afterward, where they were paid in front of each other. In the impersonal version, subjects never met, had no chance of subsequent encounter, and received no information about one another.

Given the previously reported activations in the anterior and posterior portions of the medial cingulate during a social exchange (*28*), a detailed analysis of the cingulate cortex in each pair of subjects was per-

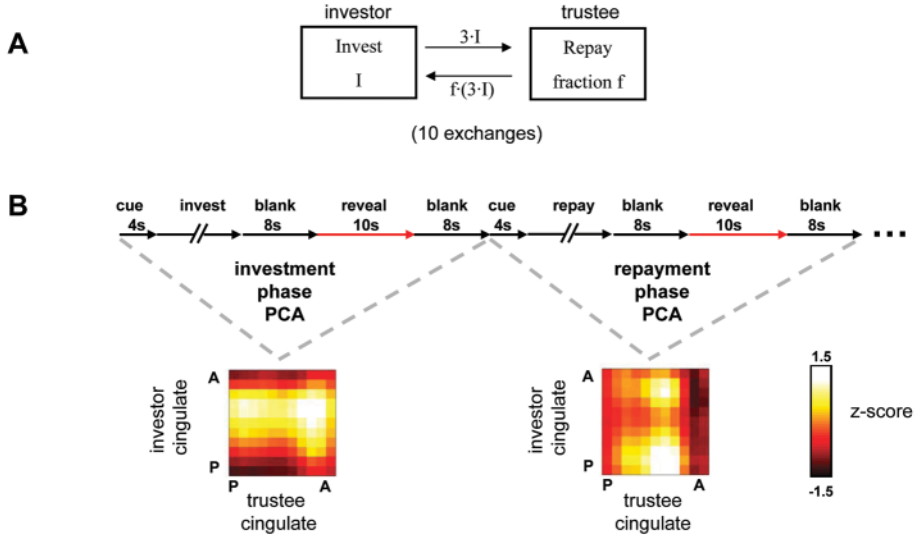


Fig. 1. Cross-cingulate correlations reveal complementary activity patterns across investment and repayment phase of game. (**A**) 10-round trust game. In each round the investor contributes some amount I between 0 monetary units and 20 monetary units, which is tripled ($3 \times I$) and sent to the trustee who then repays some fraction f of the tripled amount ($f \times 3 \times I$). (**B**) Cross-cingulate principal component analysis (PCA) revealed distinct, but complementary patterns when applied to the cross-correlations between cingulate cortices of investor and trustee (*34*). A similar result was found using independent component analysis on the same data (fig. S5).

[1]Human Neuroimaging Laboratory, Department of Neuroscience, [2]Department of Psychiatry, Center for Theoretical Neuroscience, Baylor College of Medicine, 1 Baylor Plaza, Houston, TX 77030, USA. [3]Social Cognitive Neuroscience Laboratory, California Institute of Technology, Division of Humanities and Social Sciences 228-77, California Institute of Technology, Pasadena, CA 91125, USA.

*These authors contributed equally to this work.

formed. We segmented the medial cingulate and the surrounding paracingulate cortex into separate spatial domains (*33*), computed cross-cingulate and cross-paracingulate correlation matrices for different lags in each phase of the task (investment phase and repayment phase), and carried out temporal principal component analysis (PCA) on the resulting three-dimensional correlation matrix (Fig. 1B) (*34*, *35*). Analysis yielded complementary spatial patterns for cingulate cortices (Fig. 1B)—that is, patterns of activation in one phase were transposed across role when analysis was performed for the other phase. Similar results were found using independent component analysis (ICA) on the cross-cingulate correlation matrix (fig. S5) (*33*).

The cross-cingulate analysis led us to examine the hemodynamic time series in each cingulate segment. This region-of-interest analysis revealed three distinct response types (Fig. 2A). The first followed the submission of a subject's own decision (unimodal "own"-dominated response); the second followed the visual presentation of a partner's decision (unimodal "other"-dominated response). This is a remarkable finding, because visual presentation of the subject's own decision elicited little response in the cingulate cortex. The third response type was bimodal, yielding approximately equal responses after submission of one's own decision and revelation of the partner's decision. However, the peak amplitude of these distinct response types was not uniform across the anterior-posterior axis of the cingulate. Instead, they displayed a systematic spatial variation that was complementary across the basic response types ("own" and "other"). Specifically, the submission of one's own decision elicited maximal activation in middle cingulate regions (Fig. 2A, segment G), whereas viewing the revelation of a partner's decision yielded maximal activation in anterior and posterior cingulate (an example of an anterior response is shown in Fig. 2A, segment K). This result was in stark contrast to the results of the paracingulate analysis, which indicated that, although the dorsal anterior cingulate cortex was highly activated during the experiment, there was no spatial selectivity for either stimulus. In fact, the dorsal anterior cingulate cortex responded strongly to the submission of decisions and the revelation of partner choices, and it was the only paracingulate region significantly activated by either (fig. S6) (*33*).

The distinct response types and the systematic spatial variation of peak amplitudes across the anterior-posterior axis disappeared completely in motor control (n = 15; Fig. 2C) and sensory control experiments (n = 17; Fig. 2E) not involving exchange with another agent (*33*). In the motor control, subjects reiterated the motor responses of randomly selected investors. We applied the same region-of-interest analysis to the control data (Fig. 2 and fig. S2). Statistical comparison of responses in each of the cingulate domains showed that responses differed significantly between the normal trust task and the control tasks (*33*). In particular, no significant response was present in the middle cingulate (Fig. 2C), ruling out the possibility that middle cingulate activation in the trust game was the result of motor activity produced by button tapping. In the sensory control, partner reveal screens from the trust game were viewed passively by a separate cohort of subjects (n = 17). Because partner reveal screens

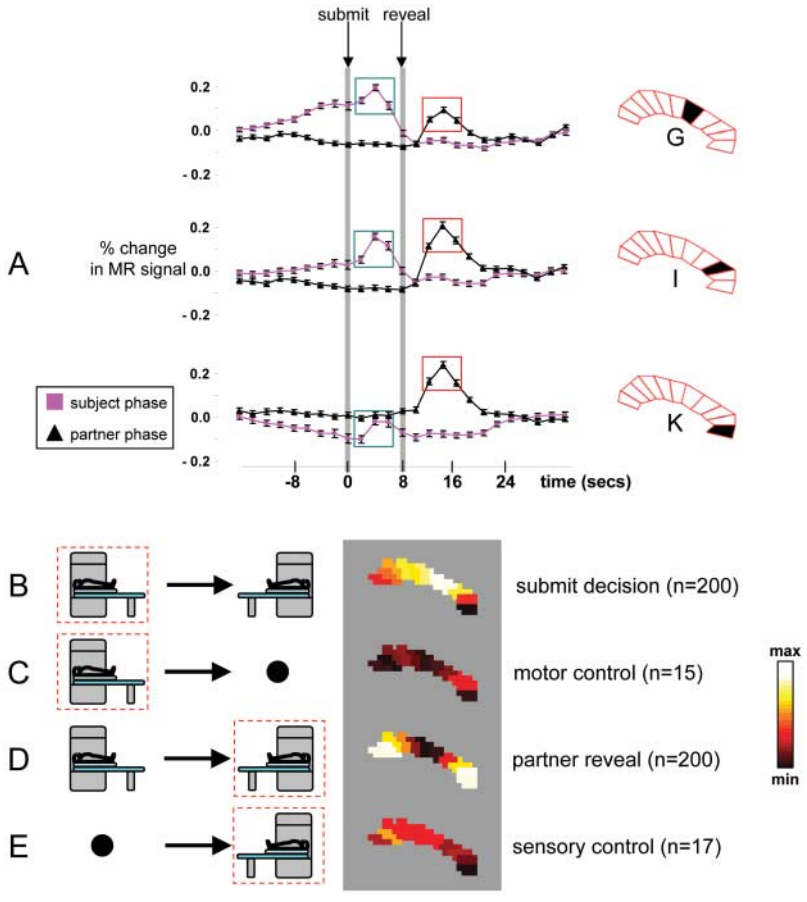


Fig. 2. Agent-specific responses and their pattern disappear outside of economic exchange. (**A**) Calculation of response pattern diagrams. Traces are the average magnetic resonance (MR) signal during subject decision phases (magenta lines) and during partner decision phases (black lines); error bars represent the standard error of the mean (n = 200 subjects). To compute the magnitude of responses to submitting a decision, MR values were selected from the time of peak response and the peak's two flanking points (teal boxes). These values, when averaged, represent the responsiveness of a segment to the submission of the decision. This measure was performed for all segments, and a pseudo color image was produced, as depicted in Figs. 2 and 3. For responses to partner reveal screens, MR values corresponding to the peak activity after screen onset and the peak's two flanking points (red boxes) were averaged and compiled into a similar pseudo color map. (**B**) The average response to submitting a decision is shown for subjects playing the linked trust game (n = 200), and a predominance of the middle cingulate is apparent. (**C**) Average response profile to submitting decisions in the unlinked motor control experiment (n = 15). No significant differentiation was observed across the cingulate of subjects in this task, but response levels in the middle cingulate were significantly different than those in the linked trust game (P = 0.00001). (**D**) Subjects from the linked trust experiment (n = 200) demonstrate the average response to viewing a social partner's decision. The predominance of responses in the anterior and posterior poles of the cingulate is apparent in this group. (**E**) Average response to viewing screens in the unlinked visual control experiment (n = 17). No significant differentiation was observed across cingulate domains, but responses in both anterior and posterior regions were significantly different than those in the linked trust game ($P < 0.01$). Maximum activation in (B) and (C) is 0.21% change in MR signal; maximum activation in (D) and (E) is 0.12%; minimum activation for each is 0.00%.

in the trust game had novel content and had been generated by an external agent, we could not use the original data set to separate responses to social or novel stimuli. Thus, subjects in the sensory control task were informed that their compensation depended on money shown under the "gave" label on the screen, but were not told about the social task from which this screen was derived (fig. S1). This manipulation was performed so that a screen's content still held novel and valuable information, but was devoid of social interaction. In each of the 11 cingulate domains, BOLD responses after each of 10 outcome screens did not resemble those obtained during the analogous presentation in the linked experiment (Fig. 2E and fig. S3). There was no systematic spatial variation in response amplitudes across the cingulate gyrus. To verify that motor control and sensory control effects were not due to the smaller number of subjects involved in the control manipulations, we confirmed this finding by using random samples of the existing 200-subject database (fig. S3).

The results provide strong support for three new findings: (i) agent-specific response types localized on the medial bank of cingulate cortex, (ii) a systematic spatial variation of each response type across the anterior-posterior axis of cingulate cortex, and (iii) a dependence of both signals on the presence of a responding agent. Despite the relative simplicity of this economic exchange game, other variable(s) related to this task may have been the underlying cause of the different response types, the spatial variation across the cingulate, and the difference in response to visual revelation of one's own decision and one's partner's decision. However, the different response types and their systematic but complementary spatial variation across the cingulate did not change as a function of a range of dimensions (Fig. 3). The most dramatic dimensions tested in Fig. 3 are reciprocity and social context (personal versus impersonal). In previously published work, reciprocity, expressed as degree of tit-for-tat behavior across rounds, acted as a powerful behavioral signal to one's partner and elicited strong, measurable neural correlates (*25*). Yet, as illustrated in Fig. 3 (bottom three rows), differences in reciprocity had no effect on the response types or on their spatial variation along the cingulate. The same result held for the difference in social context (personal, $n = 104$; impersonal, $n = 96$), where prior exposure to one's partner, the sight of their picture in each round, and the knowledge of an imminent encounter afterward had no effect. Likewise, no differences were observed when comparing subject role (investor or trustee), sex of subject, or amount of money sent or received.

Using an iterated economic exchange task, we found two distinct response types along the cingulate cortex consistent with agent-specific responses that signal "me" and "not me," Rather than residing in strictly demarcated functional zones, these complementary responses types exhibited smooth transitions across the entire medial bank of the cingulate gyrus. It is difficult to probe the extent to which a subject is considering outcomes for oneself or a social partner; individuals in a social exchange must necessarily model the actions of both agents as decisions are made and revealed. Despite this obstacle, the pattern of activation observed in these data was clearly sensitive to which participant was responsible for a given action. The response types and their variation through the tissue space disappeared in control experiments where money sent, actions taken, and money received were matched to those experienced during the normal multiround exchange (Fig. 2 and fig. S3) (*35*). These controls provide strong evidence that the response types were due to neither motor and premotor responses nor to sensory responses to outcome screens.

One question deserves separate consideration: Did the reveal screens generate simple surprise or novelty responses along cingulate that were not related to the social element of the exchange? Although this reasonable interpretation is possible, the control experiments suggest otherwise. The response pattern along the cingulate disappeared in the control experiments where subjects received stimuli that were visually identical to those in the trust game and were composed of novel, reward-related information. This manipulation used novel stimuli with economically meaningful content to probe the reveal response and showed neither an "other" response anywhere along the cingulate nor the spatial variation so prominent in the linked trust task. We take these data as strong support that the responses observed in the linked trust game were not the mere result of surprising content.

The response types and their spatial variation along the cingulate were remarkably stable to a range of manipulations. They survived the personal and impersonal treatments, did not change as a function of the reciprocity (a variable previously shown to be the major behavioral signal in this game) (*25*), and were not

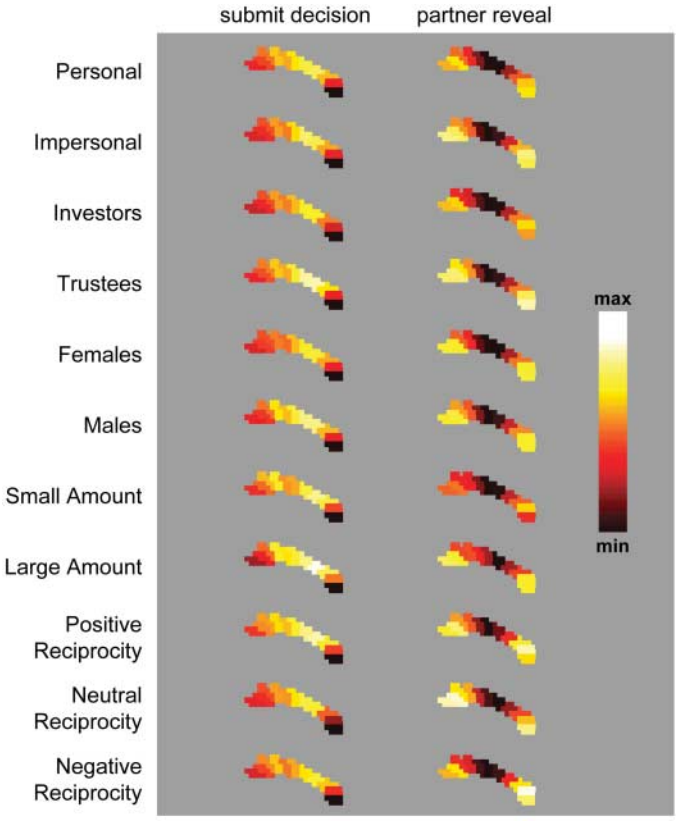


Fig. 3. Cingulate pattern of "me" and "not me" remains constant across a range of variables. All responses to decision submission are shown in the left column, whereas those responses to partner reveal screens are shown in the right column. With the exception of the reciprocity and amount diagrams, all responses were averaged across rounds before compilation. Rows labeled "Personal" and "Impersonal" separate activity across social context: the personal ($n = 104$ individuals) and impersonal ($n = 96$ individuals) tasks. Rows labeled "Investors" and "Trustees" demonstrate the consistency of the responses across the two different roles ($n = 100$ for each). Rows labeled "Males" and "Females" demonstrate that these responses do not differ across gender ($n = 100$ for each). Rows labeled "Small amount" and "Large amount" show that these patterns do not depend upon the amount of resource sent or received by the player (upper 25% versus lower 25% of payments; $n = 454$ and 161, respectively). Finally, the rows labeled "Positive," "Neutral," and "Negative" reciprocity depict responses across different valences of a behavioral variable of already known interest (*25*). These diagrams correspond to average BOLD responses to positive (values > 0.1; $n = 377$ choices), neutral ($-0.1 \leq$ values ≤ 0.1; $n = 865$ choices), and negative (values < -0.1; $n = 458$ choices) values of the reciprocity index. Left column maximum is 0.25% change in MR signal; right column maximum is 0.16%; minimum activation for both is 0.00%.

changed as a function of sex, role, amount sent, or amount received. Lastly, these signals were clear even in individual subject pairs, as shown in fig. S4 for a single interacting dyad.

The observed lack of change as a function of reciprocity is extremely important because it reduces the likelihood of two alternate interpretations of these data. The average behavior in this game is initial cooperation followed by tit-for-tat moves, a strategy conjectured to be optimal in a reciprocal interaction (*3*, *6*). To play such a tit-for-tat strategy, a player's brain must compute the expected next move of their partner and compare this to the actual outcome. Consequently, large deviations in reciprocity would also carry large prediction error signals, a signal type known to show up near or around dorsal anterior cingulate cortex (dACC) (*36*, *37*). Two possibilities arise. The error signals could activate dACC because they reflect directly an error response. Alternately, large deviations in reciprocity represent a signal with a large amount of uncertainty and might engage an output conflict response typical for this brain region (*36*–*44*). However, neither of these interpretations would anticipate an important feature of the data actually observed. There was no difference in response types or their spatial variation as a function of positive, negative, or neutral reciprocity. One would at least expect both alternate explanations to show responses that differentiated neutral reciprocity from the other two categories (positive and negative). One possibility is that our current analysis missed the error signals altogether for some unidentified reason. However, by using this same behavioral task, we have previously identified such error-related signals elsewhere in the brain and have shown these regions to be sensitive to reciprocity (*25*). Consequently, our capacity to detect these error signals elsewhere makes it less likely that we simply missed error signals in cingulate related to strong deviations in reciprocity. However, it remains a possibility that some unprobed behavioral dimension generated an error signal along cingulate cortex during this task.

In a two-person social exchange, it is crucial for each agent to know how to credit an outcome. Failure to assign this credit accurately will compromise an agent's capacity to decide on an appropriate level of cooperation with the partner—a mistake that could prove extremely costly when averaged over multiple encounters (*1*–*6*). Consequently, we suspect that these data derive from a neural mechanism dedicated to distinguishing "me" outcomes from "not me" outcomes. The systematic spatial progression of responses suggests to us that this social agency variable may be arrayed as a map; however, the current experiment cannot adequately test this provocative possibility. It is important, therefore, to note that the assignment of credit (or agency) within a social interaction necessarily implicates a variety of cognitive and emotional mechanisms. Thus, although agency parsimoniously characterizes the activations seen with these data, it may not necessarily be congruent to the underlying functions represented along the cingulate.

Extant data support a multifunctional role for the cingulate cortex, particularly in light of the extreme diversity of information that impinges on this region. Cingulate and paracingulate cortices have long been hypothesized as sites of integration of information sources that include cognitive, emotional, and interoceptive signals. Consequently, a range of functions has been ascribed to cingulate cortex (*38*–*51*), and there are disagreements over the exact variables processed and represented in these regions. However, it is reasonably clear that cingulate and paracingulate cortices contribute to normal social cognition and adaptive decision-making (*17*–*19*). The results of this paper add the important possibility that many other variables in the social domain may be arranged in such a systematic fashion through the spatial domain, a phenotype that could be disturbed in afflictions where the capacity to distinguish "me" from "not me" is impaired (*14*–*16*, *52*–*57*).

References and Notes

1. W. D. Hamilton, *J. Theor. Biol.* **7**, 1 (1964).
2. W. D. Hamilton, *J. Theor. Biol.* **7**, 17 (1964).
3. R. L. Trivers, *Q. Rev. Biol.* **46**, 35 (1971).
4. J. Maynard Smith, G. R. Price, *Nature* **146**, 15 (1973).
5. R. Axelrod, W. D. Hamilton, *Science* **211**, 1390 (1981).
6. M. A. Nowak, K. Sigmund, *Nature* **355**, 250 (1992).
7. U. Frith, C. D. Frith, *Curr. Dir. Psych. Sci.* **10**, 151 (2001).
8. W. M. Kelley *et al.*, *J. Cogn. Neurosci.* **14**, 785 (2002).
9. M. D. Lieberman, J. H. Pfeifer, in *Cognitive Neuroscience of Emotional and Social Behavior*, A. Easton, N. Emery, Eds. (Psychology Press, Philadelphia, 2005), pp. 195–235.
10. K. N. Ochsner *et al.*, *J. Cogn. Neurosci.* **16**, 1746 (2004).
11. C. A. Seger, M. Stone, J. P. Keenan, *Neuropsychologia* **42**, 1168 (2004).
12. K. Vogeley *et al.*, *Neuroimage* **14**, 170 (2001).
13. K. Vogeley, G. R. Fink, *Trends Cogn. Sci.* **7**, 38 (2003).
14. S. Baron-Cohen, M. K. Belmonte, *Annu. Rev. Neurosci.* **28**, 109 (2005).
15. C. D. Frith, U. Frith, *Science* **286**, 1692 (1999).
16. M. Brune, *Schizophr. Bull.* **31**, 21 (2005).
17. H. L. Gallagher *et al.*, *Neuropsychologia* **38**, 11 (2000).
18. E. Brunet, Y. Sarfati, M. Hardy-Bayle, J. Decety, *Neuroimage* **11**, 157 (2000).
19. B. Wicker, D. I. Perrett, S. Baron-Cohen, J. Decety, *Neuropsychologia* **41**, 139 (2003).
20. J. Decety, P. L. Jackson, J. A. Sommerville, T. Chaminade, A. N. Meltzoff, *Neuroimage* **23**, 744 (2004).
21. M. R. Delgado, R. H. Frank, E. A. Phelps, *Nat. Neurosci.* **8**, 1611 (2005).
22. D. J. de Quervain *et al.*, *Science* **305**, 1254 (2004).
23. N. I. Eisenberger, M. D. Lieberman, K. D. Williams, *Science* **302**, 290 (2003).
24. J. D. Greene, R. B. Sommerville, L. E. Nystrom, J. M. Darley, J. D. Cohen, *Science* **293**, 2105 (2001).
25. B. King-Casas *et al.*, *Science* **308**, 78 (2005).
26. K. McCabe, D. Houser, L. Ryan, V. Smith, T. Trouard, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11832 (2001).
27. J. K. Rilling *et al.*, *Neuron* **35**, 395 (2002).
28. J. K. Rilling, A. G. Sanfey, J. A. Aronson, L. E. Nystrom, J. D. Cohen, *Neuroimage* **22**, 1694 (2004).
29. A. G. Sanfey, J. K. Rilling, J. A. Aronson, L. E. Nystrom, J. D. Cohen, *Science* **300**, 1755 (2003).
30. M. Bhatt, C. Camerer, *Games Ec. Behav.* **52**, 424 (2005).
31. T. Singer *et al.*, *Nature* **439**, 466 (2006).
32. P. R. Montague *et al.*, *Neuroimage* **16**, 1159 (2002).
33. Materials and methods are available as supporting material on *Science* Online.
34. A. Cichocki, S. Amari, *Adaptive Blind Signal and Image Processing* (Wiley & Sons Inc., New York, 2003).
35. A. Hyvarinen, J. Karhunen, E. Oja, *Independent Component Analysis* (Wiley & Sons Inc., New York, 2001).
36. C. B. Holroyd, M. G. H. Coles, *Psychol. Rev.* **109**, 679 (2002).
37. C. B. Holroyd, S. Nieuwenhuis, N. Yeung, J. D. Cohen, *Neuroreport* **14**, 2481 (2003).
38. C. S. Carter *et al.*, *Science* **280**, 747 (1998).
39. H. D. Critchley *et al.*, *Brain* **126**, 2139 (2003).
40. W. J. Gehring, R. T. Knight, *Nat. Neurosci.* **3**, 516 (2000).
41. M. P. Milham, M. T. Banich, E. D. Claus, N. J. Cohen, *Neuroimage* **18**, 483 (2003).
42. D. H. Weissman, B. Giesbrecht, A. W. Song, G. R. Mangun, M. G. Woldorff, *Neuroimage* **19**, 1361 (2003).
43. M. Botvinick, L. E. Nystrom, K. Fissell, C. S. Carter, J. D. Cohen, *Nature* **402**, 179 (1999).
44. C. S. Carter *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1944 (2000).
45. A. R. Damasio *et al.*, *Nat. Neurosci.* **3**, 1049 (2000).
46. P. Rainville, G. H. Duncan, D. D. Price, B. Carrier, M. C. Bushnell, *Science* **277**, 968 (2000).
47. K. N. Ochsner *et al.*, *Neuropsychologia* **39**, 219 (2001).
48. D. D. Dougherty *et al.*, *Biol. Psychiatry* **46**, 466 (1999).
49. K. L. Phan, I. Liberzon, R. C. Welsh, J. C. Britton, S. F. Taylor, *Neuropsychopharmacology* **28**, 1344 (2003).
50. G. Bush, P. Luu, M. I. Posner, *Trends Cogn. Sci.* **4**, 215 (2000).
51. F. A. Nielsen, D. Balslev, L. K. Hansen, *Neuroimage* **27**, 520 (2005).
52. N. Georgieff, M. Jeannerod, *Conscious. Cogn.* **7**, 465 (1998).
53. M. D. Lieberman, J. M. Jarcho, A. B. Satpute, *J. Pers. Soc. Psychol.* **87**, 421 (2004).
54. K. N. Ochsner *et al.*, *J. Cogn. Neurosci.* **16**, 1746 (2004).
55. C. A. Seger, M. Stone, J. P. Keena, *Neuropsychologia* **42**, 1168 (2004).
56. L. C. Johns *et al.*, *Psychol. Med.* **31**, 705 (2001).
57. J. M. Allman, K. K. Watson, N. A. Tetreault, A. Y. Hakeem, *Trends Cogn. Sci.* **9**, 367 (2005).

58. This work was supported by the Center for Theoretical Neuroscience at Baylor College of Medicine (P.R.M.), National Institute on Drug Abuse (NIDA) grant DA11723 (P.R.M.), National Institute of Neurological Disorders and Stroke grant NS045790 (P.R.M.), National Institute of Mental Health grant MH52797 (P.R.M.), NIDA grant DA14883 (G. Berns), the Angel Williamson Imaging Center (P.R.M.), The Kane Family Foundation (P.R.M.), The David and Lucile Packard Foundation (S.R.Q. and C.F.C.), and The Gordon and Betty Moore Foundation (S.R.Q. and C.F.C.). P.R.M. also acknowledges support from the Institute for Advanced Study (Princeton, NJ) during completion of this work. We thank the Hyperscan Development Team at Baylor College of Medicine for Network Experiment Management Object (NEMO) software implementation (www.hnl.bcm.tmc.edu/nemo) and G. Berns for early discussions and efforts leading to the development of hyperscanning. We also thank C. Bracero, A. Harvey, S. Flaherty, J. McGee, K. Pfeiffer, R. Pruitt, and S. Gleason for technical assistance.

Supporting Online Material

www.sciencemag.org/cgi/content/full/312/5776/1047/DC1
Materials and Methods
Figs. S1 to S6
References

31 January 2006; accepted 24 April 2006
10.1126/science.1125596

Bruchpilot Promotes Active Zone Assembly, Ca^{2+} Channel Clustering, and Vesicle Release

Robert J. Kittel,[1*] Carolin Wichmann,[1,2*] Tobias M. Rasse,[1*] Wernher Fouquet,[1] Manuela Schmidt,[1] Andreas Schmid,[1] Dhananjay A. Wagh,[3] Christian Pawlu,[2] Robert R. Kellner,[4] Katrin I. Willig,[4] Stefan W. Hell,[4] Erich Buchner,[3] Manfred Heckmann,[2]† Stephan J. Sigrist[1,5]†

The molecular organization of presynaptic active zones during calcium influx–triggered neurotransmitter release is the focus of intense investigation. The *Drosophila* coiled-coil domain protein Bruchpilot (BRP) was observed in donut-shaped structures centered at active zones of neuromuscular synapses by using subdiffraction resolution STED (stimulated emission depletion) fluorescence microscopy. At *brp* mutant active zones, electron-dense projections (T-bars) were entirely lost, Ca^{2+} channels were reduced in density, evoked vesicle release was depressed, and short-term plasticity was altered. BRP-like proteins seem to establish proximity between Ca^{2+} channels and vesicles to allow efficient transmitter release and patterned synaptic plasticity.

Synaptic communication is mediated by the fusion of neurotransmitter-filled vesicles with the presynaptic membrane at the active zone, a process triggered by Ca^{2+} influx through clusters of voltage-gated channels (*1*, *2*). The spacing between Ca^{2+} channels and vesicles at active zones is particularly thought to influence the dynamic properties of synaptic transmission (*3*).

The larval *Drosophila* neuromuscular junction (NMJ) is frequently used as a model of glutamatergic synapses (*4*, *5*). The monoclonal antibody Nc82 specifically stains individual active zones (fig. S1A) (*6*, *7*) by recognizing a coiled-coil domain protein of roughly 200 kD named Bruchpilot (BRP) (*6*). BRP shows homologies to the mammalian active zone components CAST [cytoskeletal matrix associated with the active zone (CAZ)–associated structural protein] (*8*), also called ERC (ELKS, Rab6-interacting protein 2, and CAST) (*9*). Whereas confocal microscopy recognized diffraction limited spots, the subdiffraction resolution of stimulated emission depletion (STED) fluorescence microscopy (*10*, *11*) revealed donut-shaped BRP structures at active zones (Fig. 1A). Viewed perpendicular to the plane of synapses, both single and multiple "rings" were uncovered, of similar size to freeze-fracture-derived estimates of fly active zones (*12*) (average length of isolated rings was 0.191 ± 0.002 μm, n = 204; average length of single rings of double ring structures was 0.148 ± 0.002 μm, n = 426; average length of double rings was 0.297 ± 0.005, n = 213) (fig. S1B). The donuts were up to 0.16 μm high, as judged by images taken parallel to the synaptic plane (Fig. 1A).

BRP seemed to demark individual active zones associated with clusters of Ca^{2+} channels. Transposon-mediated mutagenesis allowed

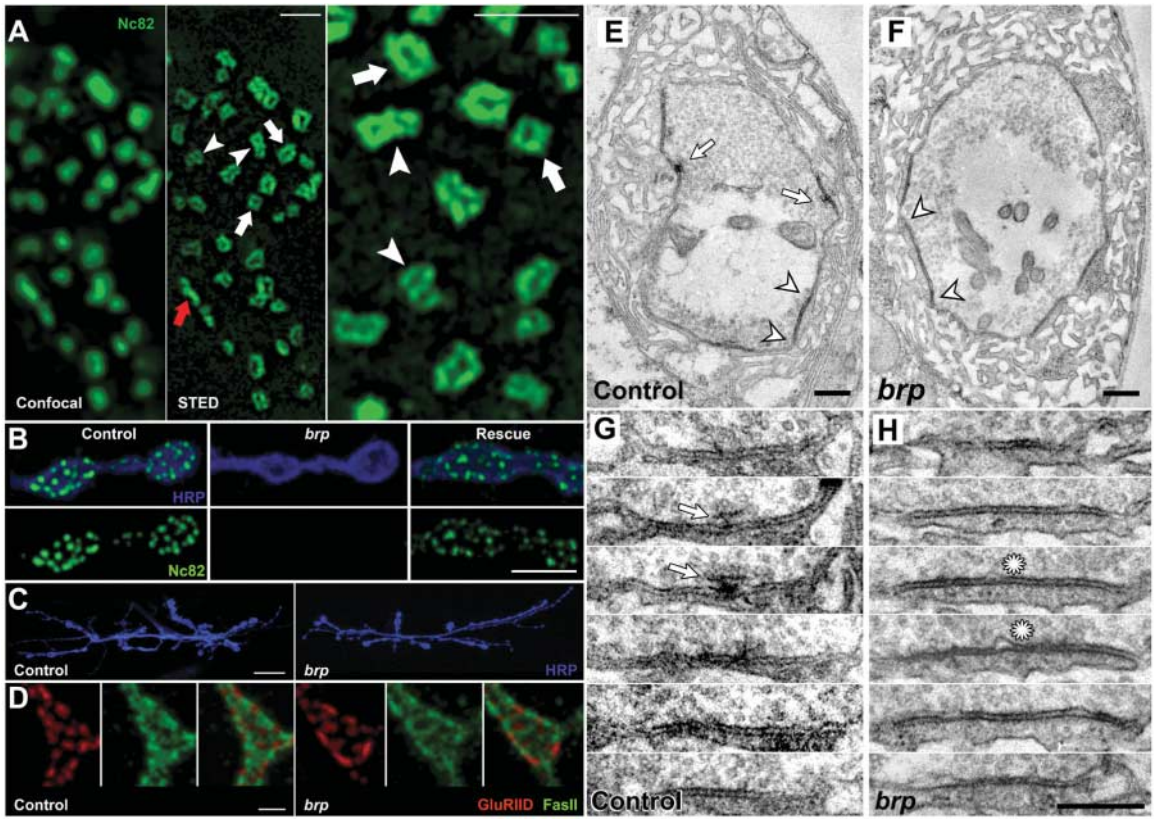


Fig. 1. Junctional and ultrastructural assembly in mutants of the active zone component BRP. (**A**) Unlike confocal, STED microscopy revealed donut-shaped structures recognized by Nc82. Viewed from above, both single (white arrows) and clusters of multiple rings (arrowheads) were identified. The red arrow indicates a synapse viewed parallel to the synaptic plane. (**B**) Individual synapses of control animals were labeled by Nc82, whereas *brp* mutant synapses completely lacked the Nc82 signal, which could be partially restored by re-expressing the *brp* cDNA in the *brp* mutant background with use of the neuron-specific driver line *ok6-GAL4*. (**C**) Staining with a neuronal membrane marker (anti–horseradish peroxidase) demonstrated normal morphological organization of *brp* NMJs. (**D**) Receptor fields were surrounded by the typical perisynaptic expression of the neuronal cell adhesion molecule (NCAM) homolog FasciclinII (FasII) in both *brp* mutants and controls. (A) to (D) are projections of confocal stacks. Scale bars represent in (A), 1 μm; (B), 4 μm; (C), 20 μm; and (D), 2 μm. (**E**) Electron micrograph of a control type Ib bouton with synapses (arrowheads) presynaptically decorated with T-bars (arrows). (**F**) A *brp* mutant bouton showing an overall normal organization but without T-bars. (**G**) Serial sections of a control synapse. A T-bar can be observed in two consecutive sections (arrows). (**H**) Serial sections of a representative *brp* mutant synapse completely lacking a T-bar and revealing presynaptic membrane rufflings (asterisks). Scale bars in (G) and (H), 250 nm.

us to isolate a mutant chromosome (brp^{69}) in which nearly the entire open reading frame of BRP was deleted (fig. S1C). *brp* mutants [brp^{69}/*df(2R)BSC29*] developed into mature larvae but did not form pupae. The Nc82 label was completely lost from the active zones of *brp* mutant NMJs but could be restored by re-expressing the *brp* cDNA (*6*) in the *brp* mutant background with use of the neuron-specific driver lines *ok6-GAL4* (Fig. 1B) or *elav-GAL4*. This also rescued larval lethality. Mutants had slightly smaller NMJs (average control size was 780.0 ± 35.8 μm², n = 14; average *brp* size was 593.3 ± 29.1 μm², n = 12; P = 0.0013) (Fig. 1C) and somewhat fewer individual synapses (average synapse number for control was 411.1 ± 41.5, n = 9; for *brp*, 296.3 ± 28.9; n = 8; P = 0.036). However, individual receptor fields, identified by the glutamate receptor subunit GluRIID (*13*), were enlarged in *brp* mutants (average field size in control was 0.43 ± 0.02 μm², n = 9; in *brp*, 0.64 ± 0.03 μm²; n = 8; P < 0.001) (Fig. 1D). Thus, principal synapse formation occurred in *brp* mutants, with individual postsynaptic receptor fields increased in size but moderately decreased in number.

In electron micrographs of *brp* mutant NMJs, synapses with pre- and postsynaptic membranes in close apposition were present at regular density (Fig. 1, F and H), and consistent with the enlarged glutamate receptor fields (Fig. 1D) postsynaptic densities appeared larger while otherwise normal (Fig. 1F). However, intermittent rufflings of the presynaptic membrane were noted (Fig. 1H), and *brp* mutants completely lacked presynaptic dense projections (T-bars). Occasionally, very little residual electron-dense material attached to the presynaptic active zone membrane was identified (fig. S2B). After re-expressing the BRP protein in the mutant background, T-bar formation could be partially restored (fig. S2C), although these structures were occasionally somewhat aberrant in shape. Thus, BRP assists in the ultrastructural assembly of the active zone and is essential for T-bar formation.

In *brp* mutant larvae we noted a drastic decrease in evoked excitatory junctional current (eEJC) amplitudes (Fig. 2A) by using two-electrode voltage clamp recordings of postsynaptic currents at low stimulation frequencies (*elav-GAL4* background control, –89.3 ± 3.4 nA; *brp*, –32.1 ± 5.9 nA; n = 10 each; P < 0.001; *ok6-GAL4* background control, –89.6 ± 4.4 nA; n = 9; *brp*, –32.8 ± 3.7 nA; n = 10; P < 0.001). This drop in current amplitude could be partially rescued through *brp* re-expression within the presynaptic motoneurons by using either *elav-GAL4* or *ok6-GAL4* (*elav-GAL4*, –55.5 ± 4.3 nA; n = 11; P = 0.01; *ok6-GAL4*, –62.2 ± 5.3 nA; n = 10; P = 0.002) (Fig. 2A). In contrast, the amplitude of miniature excitatory junctional currents (mEJCs) in response to single, spontaneous vesicle fusion events was increased over control levels (control, –0.84 ± 0.06 nA; *brp*, –1.17 ± 0.05 nA; n = 10 each; P = 0.004) (Fig. 2B). This is consistent with the enlarged individual glutamate receptor fields of *brp* mutants (Fig. 1D) and excludes a lack of postsynaptic sensitivity as the cause of the reduced eEJC amplitudes.

[1]European Neuroscience Institute Göttingen, Grisebachstrasse 5, 37077 Göttingen, Germany. [2]Institut für Klinische Neurobiologie, Josef Schneider Strasse 11, 97080 Würzburg, Germany. [3]Lehrstuhl für Genetik und Neurobiologie, Am Hubland, 97074 Würzburg, Germany. [4]Department of Nano-Biophotonics, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany. [5]Institut für Klinische Neurobiologie, Rudolf Virchow Zentrum, 97080 Würzburg, Germany.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: Heckmann_M@klinik.uni-wuerzburg.de (M.H.); ssigris@gwdg.de (S.J.S.)

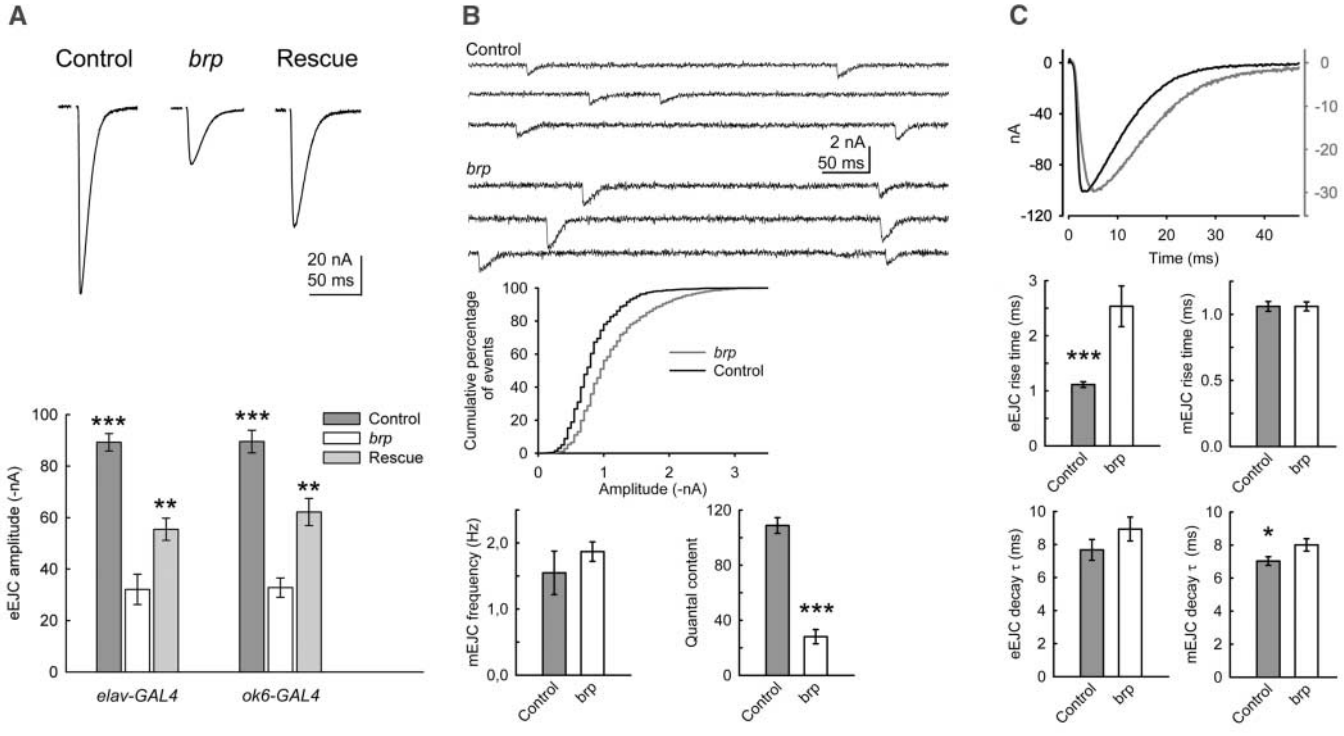


Fig. 2. Electrophysiological characterization of *brp* mutant NMJs. (**A**) (Top) Average traces of eEJCs at 0.2 Hz nerve stimulation and (bottom) mean eEJC amplitudes of control (dark gray), *brp* mutant (white), and rescued animals (light gray) carrying either a copy of *elav-GAL4* or *ok6-GAL4*. (**B**) Sample traces of mEJCs and a cumulative histogram of the amplitude distribution (0.05 nA bins). The average mEJC amplitude was increased in *brp* mutants, whereas the frequency was not significantly altered. Quantal content of *brp* NMJs was significantly reduced with respect to controls. (**C**) Average scaled eEJCs (control, black; *brp*, gray) illustrate the delayed release in *brp* mutants compared to controls. Although the rise time of eEJCs was significantly increased at *brp* NMJs, the rise time of mEJCs was indistinguishable from the control. The decay time constant (τ) of eEJCs was not significantly altered at *brp* synapses (τ control: 7.7 ± 0.6 ms; τ *brp*: 8.9 ± 0.7 ms; n = 10 each, P = 0.104), whereas mEJCs decayed with a slightly but significantly longer τ in the mutant than in the control (τ control: 7.0 ± 0.3 ms; τ *brp*: 8.0 ± 0.4 ms; n = 10 each, P = 0.045). One asterisk indicates $P \leq 0.05$, two asterisks, $P \leq 0.01$; and three asterisks, $P \leq 0.001$. Error bars indicate SEM.

It follows that the number of vesicles released per presynaptic action potential (AP) (quantal content) was severely compromised at *brp* mutant NMJs (control, 109 ± 5.7; *brp*, 28 ± 5.2; n = 10 each; $P < 0.001$) (Fig. 2B) and could not be attributed solely to the moderate decrease in synapse number. The ultrastructural defects of *brp* mutant synapses may interfere with the proper targeting of vesicles to the active zone membrane and thereby impair exocytosis. The number of vesicles directly docked to active zone membranes was slightly decreased in *brp* mutants (control average of 1.10 ± 0.13 from 51 active zones, n = 3; *brp* average of 0.87 ± 0.09 from 89 active zones, n = 4; P = 0.53). However, the amplitude distribution and sustained frequency of mEJCs (control, 1.55 ± 0.33 Hz; *brp*, 1.87 ± 0.15 Hz; n = 10 each; P = 0.186) (Fig. 2B) illustrated that *brp* mutant synapses did not appear to suffer from extrasynaptic release, as would be caused by a misalignment of vesicle fusion sites with postsynaptic receptors. Consistent with the appropriate deposition of exo- and endocytotic proteins, an apparently normal distribution of Syntaxin, Dap160, and Dynamin (fig. S3) was observed at *brp* mutant synapses.

The exact amplitude and time course of AP-triggered Ca^{2+} influx in the nerve terminal governs the amplitude and time course of vesicle release (*14*). Nerve-evoked responses of *brp* mutants were delayed (rise time of 2.53 ± 0.37 ms, n = 10) when compared with controls (rise time of 1.11 ± 0.05 ms, n = 10, $P < 0.001$), whereas in contrast mEJC rise times were unchanged (control, 1.06 ± 0.04 ms; *brp*, 1.06 ± 0.03 ms; n = 10 each) (Fig. 2C). Thus, evoked vesicle fusion events were less synchronized with the invasion of the presynaptic terminal by an AP. Spatiotemporal changes in Ca^{2+} influx have a profound effect on short-term plasticity (*15*–*17*). Whereas at 10 Hz controls (n = 18) exhibited substantial short-term depression of eEJC amplitudes, *brp* mutants (n = 15) showed strong initial facilitation before stabilizing at a slightly lower but frequency-dependent steady-state current (control at 10 Hz, –54.7 ± 3.3 nA; *brp*, –35.6 ± 3.0 nA; $P < 0.001$) (Fig. 3A). As judged by the initial facilitation at 10 Hz, neither a reduction in the number of releasable vesicles nor available release sites could fully account for the low quantal content of *brp* mutants at moderate stimulation frequencies. Furthermore, the altered short-term plasticity of *brp* mutant synapses suggested a change in the highly Ca^{2+}-dependent vesicle release probability (*18*). Paired-pulse protocols were applied to the NMJ (Fig. 3B). Closely spaced stimuli lead to a buildup of residual Ca^{2+} in the vicinity of presynaptic Ca^{2+} channels, enhancing the probability of a vesicle within this local Ca^{2+} domain to undergo fusion after the next pulse (*19*). The absence of marked facilitation at control synapses (ratio at 30-ms interval of 1.1 ± 0.03) could be explained by a depletion of release-ready vesicles (*20*). At *brp* mutant NMJs, however, the prominent facilitation at short interpulse intervals (ratio at 30-ms interval of 2.0 ± 0.13, $P < 0.001$) illustrated that the enhancement of release probability strongly outweighed the depletion of releasable vesicles. Thus, initial vesicle release probability was low, and release at *brp* synapses particularly benefited from the accumulation of intracellular Ca^{2+}.

Vesicle fusion is highly sensitive to the spacing between Ca^{2+} channels and vesicles at release sites (*3*). It has been calculated that doubling this distance from 25 to 50 nm decreases the release probability threefold (*21*), and the larger this distance, the more effective the slow synthetic Ca^{2+} buffer EGTA [ethyleneglycol-*bis*(β-aminoethyl)-*N*,*N*,*N*′,*N*′-tetraacetic acid] should become in suppressing release (*22*). Indeed, after bath application of membrane permeable EGTA-AM (tetraacetoxymethyl ester of EGTA), the reduction of evoked vesicle release was more pronounced at *brp* mutant than at control NMJs (control, 64.2 ± 13.8%; *brp*, 16.7 ± 8.8%; n = 6 each; P = 0.026) (Fig. 3C).

The Ca^{2+}-channel subunit Cacophony governs release at *Drosophila* NMJs (*23*, *24*). By using a fully functional, GFP (green fluorescent protein)–labeled variant (Cac^{GFP}) (*25*), we visualized Ca^{2+} channels in vivo (*26*). Consistently, Ca^{2+} channel expression was severely reduced over the entire NMJ and at synapses lacking BRP (Fig. 3D).

Thus, we conclude that *brp* mutants suffered from a diminished vesicle release probability

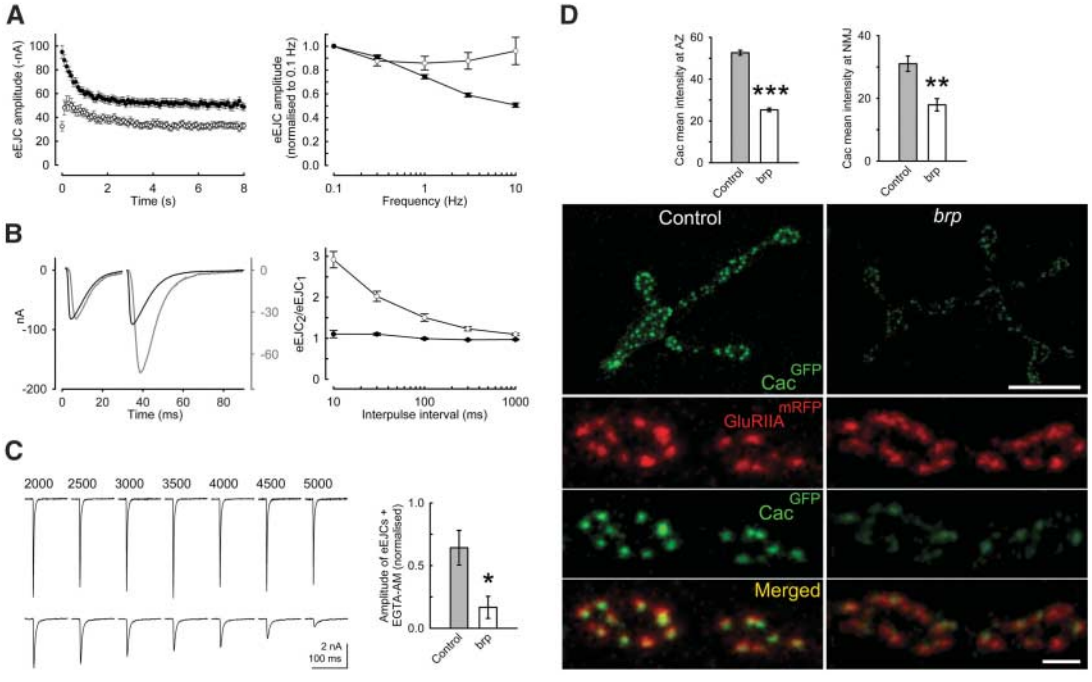


Fig. 3. Impaired vesicle release in *brp* mutants is caused by a mislocalization of presynaptic Ca^{2+} channels. (**A**) A 10-Hz stimulation revealed transient short-term facilitation of *brp* mutant currents (white circles) and the absence of a frequency-dependent depression of steady-state current amplitudes when compared with controls (black circles) ($n \geq 10$ per genotype at each frequency). (**B**) Average currents after paired-pulse stimulation at an interval of 30 ms normalized to the amplitude of the first pulse (control, black; *brp*, gray) and paired-pulse ratios at varying intervals demonstrate pronounced potentiation at *brp* NMJs (n = 9 per genotype at each interval). (**C**) Examples of nerve-evoked local postsynaptic currents recorded with a focal electrode (*36*, *37*) at indicated time points (in seconds) after bath application of EGTA-AM. The bar chart illustrates the severe reduction of current amplitudes in *brp* mutants 5000 s after EGTA-AM wash-in. The values are normalized to the initial eEJC amplitude. (**D**) Projections of confocal stacks displaying the NMJ (top images; scale bar, 10 μm) and several boutons (lower images; scale bar, 2 μm) reveal weak Cac^{GFP} signal at *brp* mutant synapses. Quantification of Cac^{GFP} intensity averaged over the entire NMJs [control, 31.1 ± 2.4 arbitrary units (a.u.); n = 13; *brp*, 18.0 ± 2.0 a.u.; n = 10; P = 0.0017] or only synaptic areas (control, 52.6 ± 1.2 a.u.; n = 421 synapses; *brp*, 25.3 ± 0.8 a.u.; n = 320 synapses; $P < 0.001$, student t test) included as bar charts. One asterisk indicates $P \leq 0.05$; two asterisks, $P \leq 0.01$; and three asterisks, $P \leq 0.001$. Error bars indicate SEM.

due to a decrease in the density of presynaptic Ca^{2+} channel clusters. It is conceivable that BRP tightly surrounds but is not part of the T-bar structure, contained within the unlabeled center of donuts. BRP may establish a matrix, required for both T-bar assembly as well as the appropriate localization of active zone components including Ca^{2+} channels, possibly by mediating their integration into a restricted number of active zone slots (*27*). Related mechanisms might underlie functional impairments of mammalian central synapses lacking active zone components (*28*) and natural physiological differences between synapse types (*17*). Electron microscopy has identified regular arrangements at active zones of mammalian CNS (central nervous system) synapses ("particle web") (*29*) and frog NMJs ("ribs") (*30*), where these structures have also been proposed to organize Ca^{2+} channel clustering. At calyx of Held synapses, both a fast and a slow component of exocytosis have been described (*31*). The fast component recovers slowly and is believed to owe its properties to vesicles attached to a matrix tightly associated with Ca^{2+} channels (*32*), whereas the slow component recovers faster (*31*) and is thought to be important for sustaining vesicle release during tetanic stimulation. In agreement with this concept, the absence or impairment of such a matrix at *brp* synapses has a profound effect on vesicle release at low stimulation frequencies, but this effect subsides as the frequency increases (Fig. 3A). The sustained frequency of mEJCs at *brp* synapses could be explained if spontaneous fusion events arise from the slow release component (*33*) or a pathway independent of evoked vesicle fusion (*34*).

Synapses lacking BRP and T-bars exhibited a defective coupling of Ca^{2+} influx with vesicle fusion, whereas the vesicle availability did not appear rate-limiting under low frequency stimulation. The activity-induced addition of presynaptic dense bodies has been proposed to elevate vesicle release probability (*35*). Our work supports this hypothesis and suggests an involvement of BRP or related factors in synaptic plasticity by promoting Ca^{2+} channel clustering at the active zone membrane.

References and Notes

1. R. G. Zhai, H. J. Bellen, *Physiology (Bethesda)* **19**, 262 (2004).
2. B. Katz, R. Miledi, *Proc. R. Soc. London Ser. B* **161**, 496 (1965).
3. E. Neher, *Neuron* **20**, 389 (1998).
4. H. L. Atwood, S. Karunanithi, *Nat. Rev. Neurosci.* **3**, 497 (2002).
5. Y. H. Koh, L. S. Gramates, V. Budnik, *Microsc. Res. Tech.* **49**, 14 (2000).
6. D. A. Wagh *et al.*, *Neuron* **49**, 833 (2006).
7. T. Wucherpfennig, M. Wilsch-Brauninger, M. Gonzalez-Gaitan, *J. Cell Biol.* **161**, 609 (2003).
8. T. Ohtsuka *et al.*, *J. Cell Biol.* **158**, 577 (2002).
9. Y. Wang, X. Liu, T. Biederer, T. C. Sudhof, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 14464 (2002).
10. T. A. Klar, S. Jakobs, M. Dyba, A. Egner, S. W. Hell, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8206 (2000).
11. S. W. Hell, *Nat. Biotechnol.* **21**, 1347 (2003).
12. C. J. Feeney, S. Karunanithi, J. Pearce, C. K. Govind, H. L. Atwood, *J. Comp. Neurol.* **402**, 197 (1998).
13. G. Qin *et al.*, *J. Neurosci.* **25**, 3209 (2005).
14. E. F. Barrett, C. F. Stevens, *J. Physiol.* **227**, 691 (1972).
15. R. S. Zucker, W. G. Regehr, *Annu. Rev. Physiol.* **64**, 355 (2002).
16. H. L. Atwood, *Nature* **215**, 57 (1967).
17. A. Rozov, N. Burnashev, B. Sakmann, E. Neher, *J. Physiol.* **531**, 807 (2001).
18. J. S. Dittman, A. C. Kreitzer, W. G. Regehr, *J. Neurosci.* **20**, 1374 (2000).
19. B. Katz, R. Miledi, *J. Physiol.* **195**, 481 (1968).
20. H. von Gersdorff, R. Schneggenburger, S. Weis, E. Neher, *J. Neurosci.* **17**, 8137 (1997).
21. M. R. Bennett, L. Farnell, W. G. Gibson, *Biophys. J.* **78**, 2201 (2000).
22. E. M. Adler, G. J. Augustine, S. N. Duffy, M. P. Charlton, *J. Neurosci.* **11**, 1496 (1991).
23. F. Kawasaki, R. Felling, R. W. Ordway, *J. Neurosci.* **20**, 4885 (2000).
24. H. Kuromi, A. Honda, Y. Kidokoro, *Neuron* **41**, 101 (2004).
25. F. Kawasaki, B. Zou, X. Xu, R. W. Ordway, *J. Neurosci.* **24**, 282 (2004).
26. T. M. Rasse *et al.*, *Nat. Neurosci.* **8**, 898 (2005).
27. Y. Q. Cao *et al.*, *Neuron* **43**, 387 (2004).
28. W. D. Altrock *et al.*, *Neuron* **37**, 787 (2003).
29. G. R. Phillips *et al.*, *Neuron* **32**, 63 (2001).
30. M. L. Harlow, D. Ress, A. Stoschek, R. M. Marshall, U. J. McMahan, *Nature* **409**, 479 (2001).
31. T. Sakaba, E. Neher, *Neuron* **32**, 1119 (2001).
32. T. Sakaba, A. Stein, R. Jahn, E. Neher, *Science* **309**, 491 (2005).
33. J. Trommershauser, R. Schneggenburger, A. Zippelius, E. Neher, *Biophys. J.* **84**, 1563 (2003).
34. Y. Sara, T. Virmani, F. Deak, X. Liu, E. T. Kavalali, *Neuron* **45**, 563 (2005).
35. J. M. Wojtowicz, L. Marin, H. L. Atwood, *J. Neurosci.* **14**, 3688 (1994).
36. J. Dudel, *Pflugers Arch.* **391**, 35 (1981).
37. C. Pawlu, A. DiAntonio, M. Heckmann, *Neuron* **42**, 607 (2004).

38. We thank E. Neher for comments on the manuscript and C. Quentin for technical support. S.J.S. (SI 849/2-1, SFB406/A16, and Research Center for Molecular Physiology of the Brain Göttingen), M.H. (HE 2621/4-1), and E.B. and D.A.W. (SFB581/B6 and GRK200/3) were supported by the Deutsche Forschungsgemeinschaft; S.W.H., by the German Ministry of Research (BMBF); and T.M.R., by a Max Planck Fellowship. The European Neuroscience Institute Göttingen (ENI-G) is jointly funded by the Göttingen University Medical School, the Max Planck Society, and Schering AG.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1126308/DC1
Materials and Methods
Figs. S1 to S3
References

16 February 2006; accepted 4 April 2006
Published online 13 April 2006;
10.1126/science.1126308
Include this information when citing this paper.

A Systems Approach to Mapping DNA Damage Response Pathways

Christopher T. Workman,[1]* H. Craig Mak,[1]* Scott McCuine,[1] Jean-Bosco Tagne,[2] Maya Agarwal,[1] Owen Ozier,[2] Thomas J. Begley,[3] Leona D. Samson,[4] Trey Ideker[1]†

Failure of cells to respond to DNA damage is a primary event associated with mutagenesis and environmental toxicity. To map the transcriptional network controlling the damage response, we measured genomewide binding locations for 30 damage-related transcription factors (TFs) after exposure of yeast to methyl-methanesulfonate (MMS). The resulting 5272 TF-target interactions revealed extensive changes in the pattern of promoter binding and identified damage-specific binding motifs. As systematic functional validation, we identified interactions for which the target changed expression in wild-type cells in response to MMS but was nonresponsive in cells lacking the TF. Validated interactions were assembled into causal pathway models that provide global hypotheses of how signaling, transcription, and phenotype are integrated after damage.

Exposure of cells to chemical and physical damaging agents can result in DNA lesions that contribute to the onset of cancer, aging, immune deficiencies, and other degenerative diseases (*1*). DNA damage is sensed by a highly conserved mechanism involving the ATM/ATR protein kinases in humans (ataxia-telangiectasia mutated/ataxia-telangiectasia and Rad3-related; homologous to Tel1 and Mec1 in yeast). These aggregate at DNA lesions (*2*) and activate signaling cascades that include the Chk protein kinases (Chk1, Rad53, and Dun1 in yeast). Chk kinases, in turn, trigger both transcriptional and transcription-independent responses, including activation of DNA repair machinery and cell-cycle arrest (*1*).

Beyond the known DNA repair genes, genomewide expression profiling in yeast has identified several hundred genes (*3*–*5*) whose expression is increased or decreased in response to alkylation damage by methyl-methanesulfonate (MMS). At the level of growth phenotype, sys-

[1]University of California San Diego, La Jolla, CA 92093, USA. [2]Whitehead Institute for Biomedical Research, Cambridge, MA 02139, USA. [3]University of Albany–State University at New York, Rensselaer, NY 12144, USA. [4]Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: trey@bioeng.ucsd.edu

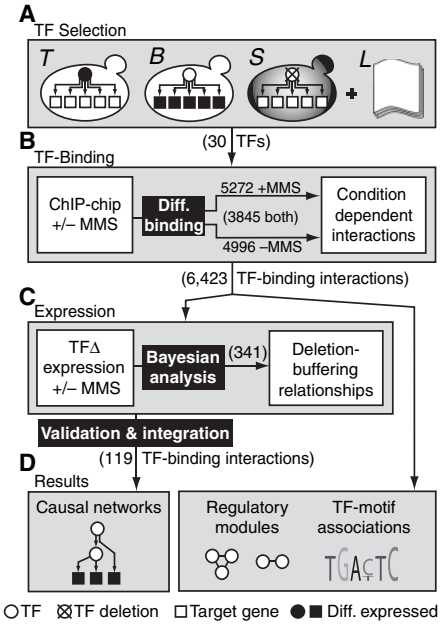


Fig. 1. Overview of the systems approach (see text).

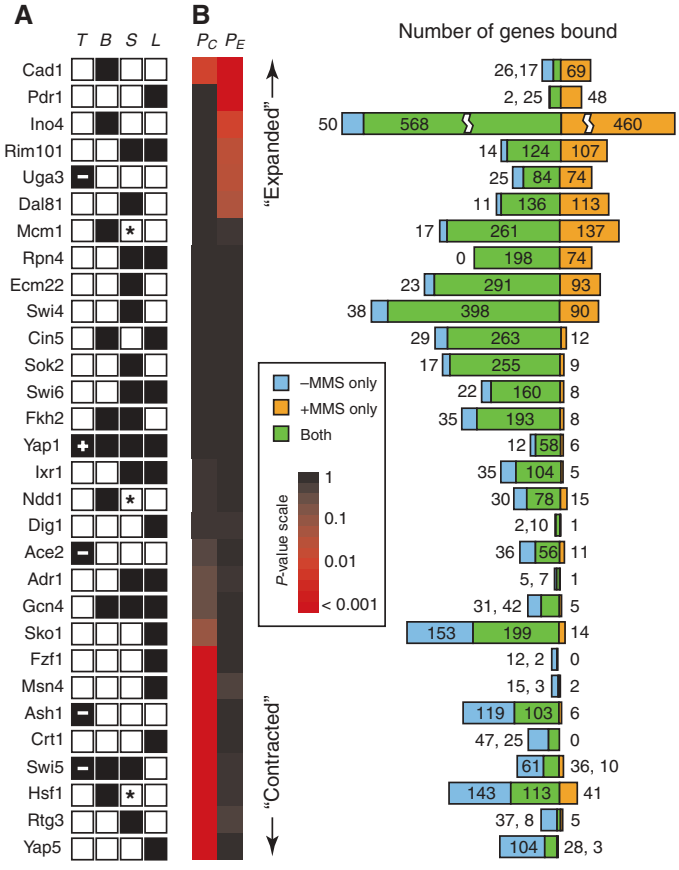


Fig. 2. TF selection and ChIP-chip experiments. (**A**) Results of the four criteria used to select the 30 TFs [TF (T) expression, expression of bound (B) genes, sensitivity (S), or literature (L); see text]. In column T, a "+" or "−" represents increased or decreased expression, respectively. For column S, a "*" denotes essential TFs. (**B**) Number of gene promoters bound by each TF. The three regions represent promoters bound exclusively in the absence of MMS (blue), presence of MMS (orange), or in both conditions (green). The proportion of genes in each region was compared to a negative control data set using Fisher's exact test (*16*). A red square in the left column (P_C, "Contracted") indicates that the proportion absence/both is significantly higher than expected in the negative control. Similarly, red in the right column (P_E, "Expanded") indicates that the proportion presence/both is higher.

tematic deletion studies have also identified several hundred genes that are required for normal recovery from alkylation damage (*6*–*8*). Surprisingly, the set of genes that, when deleted, affect damage recovery is not enriched for genes whose transcript levels change upon damage exposure (*7*, *9*). Thus, neither transcriptional profiling alone, nor genomic phenotyping alone, adequately defines the cellular response to DNA-damaging agents. However, these studies do suggest that the DNA damage response involves multiple levels of regulation, affecting not only DNA repair genes but also genes that influence protein and lipid turnover, cytoskeleton remodeling, and general stress pathways.

To construct a global model of yeast transcriptional networks activated by MMS, we applied a systems approach (*10*) that integrated data from genomewide chromatin immunoprecipitation (ChIP) assays, expression profiling, systematic phenotyping, and protein interaction databases (Fig. 1). First, we performed a systematic screen for transcription factors (TFs) involved in the MMS response. TFs were chosen from a set of 141 yeast DNA binding factors (*11*) and were selected according to any one of three criteria (Fig. 2A). These criteria were TF expression (that is, the TF was differentially expressed after exposure to 0.03% MMS); expression of bound genes [that is, the TF had been previously shown (*11*) to bind the promoters of genes that were differentially expressed in the above MMS experiment]; or sensitivity [that is, deletion of the TF gene, if not lethal, caused growth sensitivity in MMS relative to that in nominal conditions (fig. S1)].

A set of 23 TFs was identified (Fig. 2A). Four TFs were implicated by multiple criteria (Yap1, Gcn4, Fkh2, Swi5), and three were encoded by

Table 1. DNA sequence motifs found in promoters bound in only one condition.

Consensus sequence*	Promoter set†	No. of promoters with motif (−MMS)‡	No. of promoters with motif (+MMS)‡	Motif source§
	Motifs found in −MMS but not +MMS‖			
TGACTC	Gcn4	20/31	0/5	GCN4¶, BAS1#
ATTAGTAAGC	Cad1	18/26	0/69	CAD1#
cGGGGG	Hsf1	35/143	1/41	ADR1#
CGGGGCACnCTcStCCG	Hsf1	43/143	3/41	GAL4¶, PUT3#
TTCtannnnnTTC	Hsf1	44/143	7/41	HSF1#
CCGGtACCGG	Hsf1	37/143	0/41	LEU3#
CGGCGCCCGCGAn	Hsf1	43/143	6/41	RFA2
GCCSnGSCC	Hsf1	40/143	1/41	SKN7#
GCGGCnnnGCGGC	Hsf1	39/143	2/41	STP1#
ACCCGTACAt	Yap5	44/104	0/3	SFP1#
	*Motifs found in +MMS but not −MMS***			
TMSCTGCAAAnTT	Gcn4	0/31	5/5	Predicted (ANN-Spec)
GCTCGAAAA	Ndd1	3/30	12/15	Predicted (ANN-Spec)
TGAYTAACn	Sko1	7/153	11/14	Predicted (ANN-Spec)
TnTCnCTCAT	Swi5	6/61	9/10	Predicted (ANN-Spec)
GCGGCnnnGCGGC	Pdr1	0/2	21/48	STP1#
GCSGGGnCGG	Pdr1	0/2	19/48	SUT1#

*Letter based on IUPAC code; lowercase used when information content < 0.3 bits per position. †The set of promoters bound by each TF was analyzed for the presence of known and novel motifs. ‡Motif presence was scored using ANN-Spec at a threshold that predicted sites at a rate of $<10^{-4}$ over all intergenic regions. §Predicted motifs were identified using ANN-Spec. Known motifs were drawn from the SCPD database (¶) or from Harbison *et al.* (*17*) (#). ‖Enriched in −MMS-only promoters (hypergeometric $P \leq 10^{-7}$) but not in +MMS-only ($P > 10^{-2}$). **Enriched in +MMS-only promoters (hypergeometric $P \leq 10^{-7}$) but not in −MMS-only ($P > 10^{-2}$).

essential genes (Hsf1, Mcm1, Ndd1). Two of the 14 MMS-sensitive TF-deletion strains (*sok2Δ* and *ecm22Δ*) had not been reported in a previous genomewide assay for MMS sensitivity (*6*). The set also included seven of the nine known cell-cycle regulators (*12*), as might be expected given that DNA damage affects cell-cycle progression (*1*). A search of the literature identified 15 TFs that had been previously associated with regulation of DNA repair (*13*–*15*). Eight of these were also detected by our systematic criteria; the remaining seven were added to our list to yield a total of 30 TFs associated with the DNA damage response.

We then used the technique of chromatin immunoprecipitation coupled with microarray chip hybridization (ChIP-chip) to identify the MMS-induced transcriptional network immediately downstream of each TF (Fig. 2B). Exponentially growing yeast cultures were exposed to 0.03% MMS for 1 hour, resulting in ~50% cell viability (*3*, *6*). At a significance threshold of $P \leq 0.001$, a total of 5272 protein-DNA interactions with 2599 distinct genes were identified for the 30 factors. To reveal how the transcriptional network was reprogrammed between damaging and nondamaging conditions, we compared the MMS-induced interactions for each TF to the corresponding interactions as reported in nominal (noninduced) growth conditions by Lee *et al*. (*11*). Raw data from nominal conditions were reanalyzed using a data processing pipeline and significance threshold ($P \leq 0.001$) that were identical to those used for the MMS experiments, yielding a total of 4996 protein-DNA interactions with 2588 distinct genes for the same 30 factors in nominal conditions. The detailed experimental methods and raw data files (tables S1 to S4) are provided in the supporting online material.

For each TF, we applied a pairwise statistical analysis to score the overlap in promoter binding between the damage-induced and noninduced conditions (*16*). This method exploits the dependency in *P* values of binding between two related ChIP-chip data sets to identify TF-promoter interactions with more sensitivity than can be obtained if each data set were analyzed separately (fig. S2). Using this method, the number of promoters bound by each TF ranged from 13 (Adr1 and Dig1) to 1078 (Ino4) with an average of 214 per factor. Six factors bound significantly ($P < 0.05$) more genes in the presence of MMS, and conversely, eight factors bound significantly more genes under nominal growth conditions (Fig. 2B).

Several promoter sets were enriched for DNA sequence binding motifs reported in the literature (Table 1) (*17*–*19*). The observed shifts in promoter binding for Cad1 and Hsf1 correlated with the known Cad1 and Hsf1 binding motifs, which were enriched in the sets of promoters bound before but not after MMS exposure. We also used the ANN-Spec algorithm (*16*, *20*) to search each set of promoters for motifs that had not been previously reported. Four such motifs were found, associated with Gcn4, Sko1, Ndd1, or Swi5, respectively (Table 1). The GCTCGAAAA motif was found upstream of 12 of 15 genes bound by Ndd1 in the presence of MMS, but was found upstream of only 3 of 30 genes bound in the absence of MMS. Such motifs may have been missed in past analyses because they are not active before damage exposure. They may represent DNA binding sequences bound directly by the associated factor (Gcn4, Sko1, Ndd1, or Swi5) upon posttranslational modification, or alternatively, they may be bound by a different TF that is coordinately recruited to the promoters after MMS treatment.

To identify combinations of TFs that regulate genes in common, we scored the significance of overlap between the gene sets bound by each pair of the 30 TFs (Fig. 3A) (hypergeometric test at $P \leq 0.01$). Cad1 was found to pair with Yap1 in the absence of MMS but paired with Hsf1 in the presence of MMS. Several cell-cycle TFs that normally co-regulate large numbers of genes (such as Fkh2 and Swi6 or Ace2 and Swi5) (*12*) no longer appeared to do so after MMS treatment, perhaps because DNA damage causes delayed progression through the cell cycle. A prominent combination that emerged after MMS treatment consisted of Ino4 and six other factors (Dal81, Mcm1, Rim101, Ecm22, Rpn4, and Uga3) from the "expanded" set, which regulated several hundred genes in common (Fig. 3B). An additional set of >200 genes were targeted uniquely by Ino4 both before and after damage exposure (Fig. 3B), supporting a previous hypothesis that Ino4 is a global regulator of gene expression (*21*).

Next, we sought to validate transcriptional effects of the measured binding interactions and to pinpoint the particular interaction pathways involved in transmission of the damage response signal. For this purpose, we used yeast genome microarrays to monitor MMS-induced gene expression changes across the viable knockout strains (*22*) for the TFs found to be important in the DNA damage response (*16*) (table S5). Of the 30 TFs identified, 27 were nonessential (Fig. 2A, column S) and could be profiled. Hierarchical clustering over all genes confirmed that these knockout profiles were globally more similar to the responses of wild-type cells to MMS [as measured in this study and previously (*3*, *4*)] than to the expression responses of wild-type cells to other stress conditions (*5*, *23*)

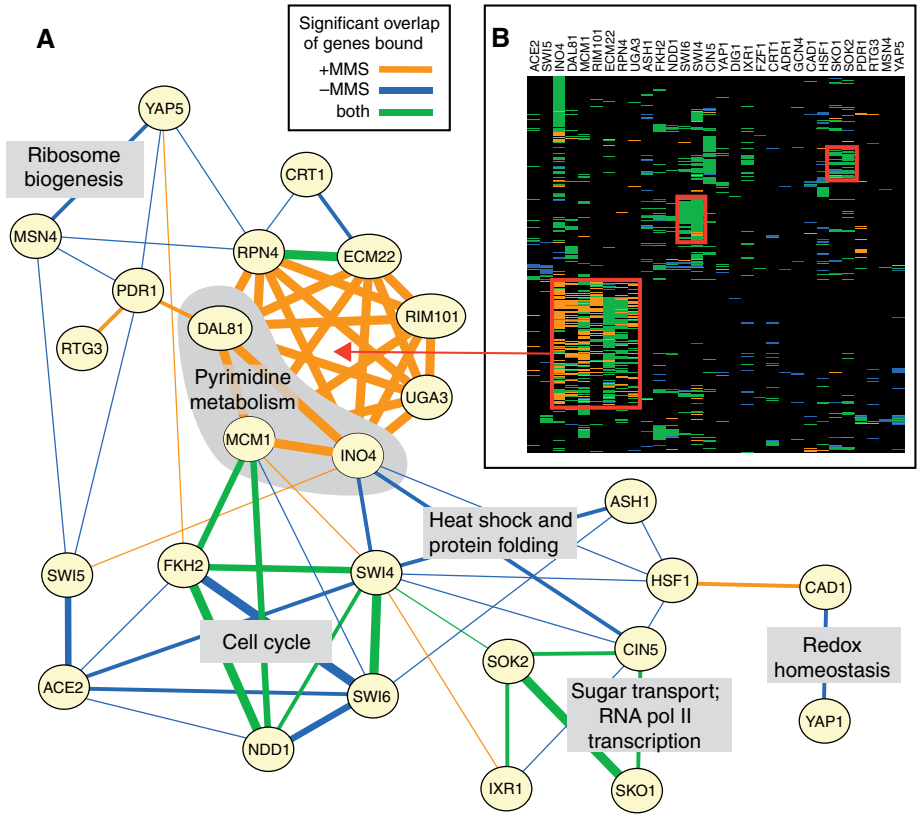


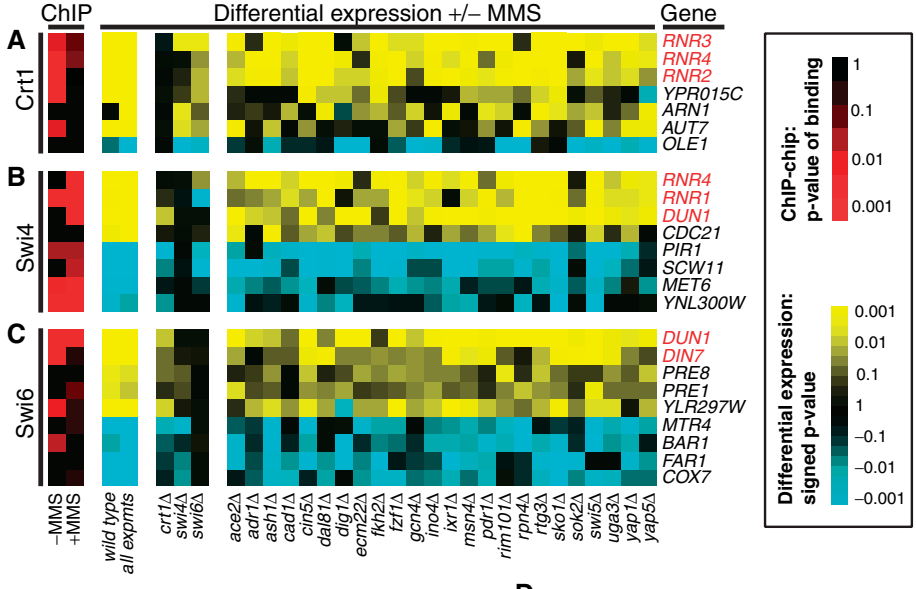
Fig. 3. Overlap in TF binding patterns in response to MMS. (**A**) TFs are linked if they bind sets of promoters that significantly overlap. Thicker lines represent overlap that is more statistically significant than overlap represented by thin lines. Linked TFs are labeled with the high-level functions of their associated sets of (co)-regulated genes. (**B**) Hierarchical clustering matrix for the 30 TFs (columns) across the top 1000 gene targets with highest variance (in log *P*-value of binding; rows) pooled over –MMS and +MMS conditions. Color is based on the binding condition(s), as in (A), or is black when the target is not bound in either condition.

(fig. S3). Furthermore, although ~20% of the transcriptional response to MMS may be due to slowed cell cycle progression through G_1/S phase (*4*, *5*), most responsive genes were not periodically expressed during the cell cycle (fig. S4).

Processed expression data were analyzed to identify genes that were genetically "buffered" by one or more TF deletions. In this context, we define "deletion buffering" to mean an effect in which genes that are normally differentially expressed become unresponsive in a specific knockout background. For each gene-TF combination, wild-type and knockout profiles were analyzed to score the significance of the deletion-buffering effect using a Bayesian scoring scheme (fig. S5, table S6) (*16*). At $P \leq 0.005$, a total of 341 genes showed deletion buffering in the 27 knockouts, corresponding to 27 genes on average over a range of 90 genes for the Adr1 knockout to 4 genes for the Ecm22 knockout (fig. S6).

As a positive control, we examined the deletion-buffering results for Crt1 (Rfx1), a transcriptional repressor of the ribonucleotide-diphosphate reductase (*RNR*) complex that catalyzes synthesis of new nucleotides during DNA repair (*24*, *25*). As expected, the expression levels of *RNR2*, *3*, and *4* were deletion-buffered in the *crt1*Δ strain but not in most other strains, and Crt1 bound the promoters of these genes before but not after MMS treatment (Fig. 4A; fig. S7). Many of the remaining deletion-buffering events represent previously undocumented regulatory relationships. For example, contrary to a previous report (*26*), we found that both members of the Swi4-Swi6 complex could bind the *DUN1* promoter and were required for the *DUN1* transcriptional response (Fig. 4, B and C).

Beyond validation of individual interactions, the deletion-buffering analysis provided insights into the damage response system as a whole. Paradoxically, the set of genes that are differentially expressed in response to DNA damage does not significantly overlap with the set of genes required for growth under damaging conditions (*7*, *9*). Our new expression data confirmed these findings for MMS, but also showed that the number of genes buffered by a TF knockout was highly correlated with its degree of MMS sensitivity ($r = 0.72$) (Fig. 4D). For example, *adr1*Δ, the most sensitive TF knockout in our study, also buffered the largest number of genes. Thus, the TFs most essential for cellular recovery after MMS exposure are, apparently, also the most central to the MMS transcriptional response.



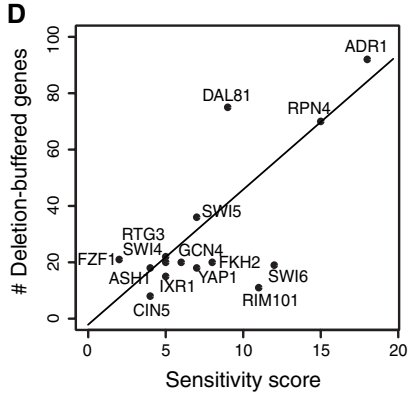


Fig. 4. Representative genes showing deletion buffering for (**A**) Crt1, (**B**) Swi4, or (**C**) Swi6. Expression changes are colored yellow for up-regulation or blue for down-regulation. Genes highlighted in red are those previously known to function in the DNA damage response. (**D**) Correlation between phenotypic sensitivity of each deletion strain in MMS (*x* axis) versus the number of genes buffered by the TF deletion (*y* axis). Sensitivity scores were drawn from (*6*). Scores range from 0 to 30, where high scores indicate that a TF deletion was found to be sensitive to low concentrations (less than 0.03%) of MMS over replicate growth experiments. Only TFs with detectable sensitivity (score > 0) are shown.

The opposite of deletion buffering is deletion enhancement, that is, genes that are MMS responsive in a TF deletion strain but not in the wild type. Deletion enhancement was a much rarer event than deletion buffering. At the same *P*-value threshold, only 16 genes showed deletion enhancement, whereas 341 showed deletion buffering (table S7). TFs associated with deletion enhancement appear to be required to maintain stable expression of a set of genes, which become MMS responsive in their absence.

Only 11% of the observed deletion-buffering events (37 out of 341) coincided with a direct ChIP-chip binding interaction (fig. S8). Such low overlap might occur for two reasons. First, failure to detect deletion buffering does not necessarily invalidate TF-promoter binding. Second, an observed deletion-buffering effect might be indirect, that is, mediated by longer regulatory pathways connecting the deleted TF to its regulated gene through one or more intermediate factors.

To identify these longer pathways, we applied a Bayesian modeling procedure (*27*) to search the known physical network for the smallest set of paths (of two interactions) that were supported by the largest number of deletion-buffering events (*16*). Types of regulatory pathways identified are shown in Fig. 5A. The physical network consisted of TF-promoter binding interactions from (i) the 5272 interactions measured in the presence of MMS for the 30 TFs in our study; (ii) the 4996 interactions measured for these TFs under nominal conditions; and (iii) 5903 interactions for the 74 additional TFs assayed by Lee *et al.* (*11*), also in nominal conditions. We also included (iv) a set of 14,319 high-throughput protein-protein interaction measurements from the Database of Interacting Proteins (*28*). These protein-protein interactions were measured in cells grown in the absence of MMS and hence might or might not be present in MMS-treated cells. In the combined network, a link from protein A to B represented the observation that protein A directly targets the promoter of the gene encoding protein B [sources (i) through (iii)] or that proteins A and B physically interact [source (iv)].

In total, we identified 68 buffering events that validated 88 longer paths. These paths were combined with the 37 direct effects identified earlier to formulate a model of the transcriptional response to MMS (Fig. 5B; tables S8 and S9). This model explains the MMS expression response of 82 genes and provides the basic scaffold on which interprocess communication is achieved in the transcriptional response to an alkylating agent. At the core of the model are the known damage response genes *RNR1*, *RNR2*, *RNR4*, *RFA1*, *RFA2*, *DIN7*, *DUN1*, and *MAG1* (*3*). Although some of the TFs regulating these genes are expected on the basis of previous studies (e.g., Swi4, Swi6, or

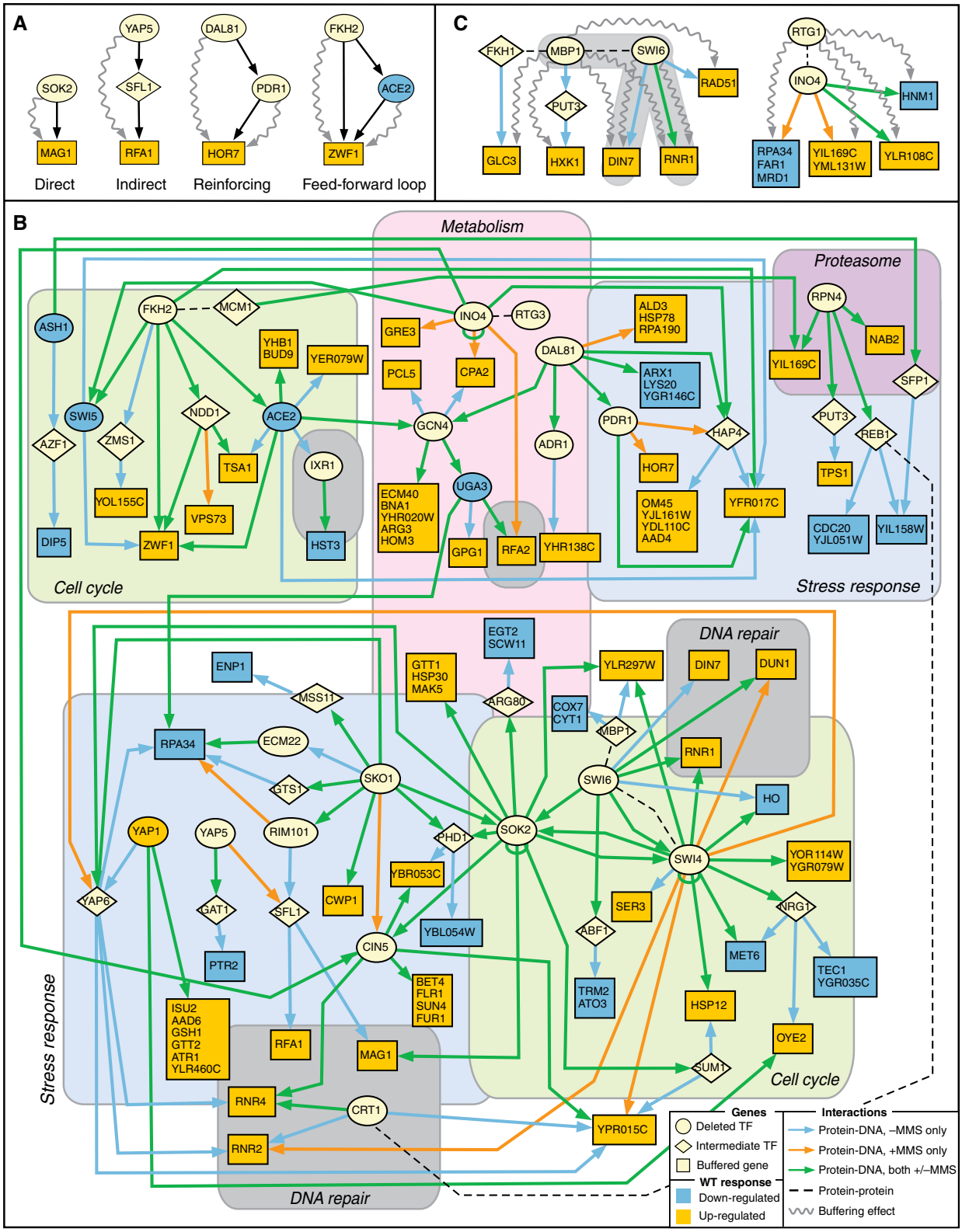


Fig. 5. A network of regulatory interactions connecting TFs to deletion-buffered genes. (**A**) Example regulatory paths in which deletion-buffering effects (squiggled arrows) support binding interactions (straight arrows). (**B**) The full validated model based on overlap between binding and buffering (see text). Buffering effects are omitted for clarity (listed in table S8). Regions of the network are organized on the basis of known functions of the deleted TFs. (**C**) Models resulting from follow-up experiments to generate buffering profiles for Mbp1 and Rtg1.

Crt1), many others would not have been predicted, including those that have been previously associated with lipid metabolism (Ino4), stress response (e.g., Yap5), or cyclic adenosine 5′-monophosphate–dependent signal transduction (Sok2). Overall, the model highlights extensive regulatory cross-talk among the processes of DNA replication and repair, cell cycle and cell-cycle arrest, stress responses, and metabolic pathways.

Every path of length two implicates an intermediate factor that is expected to regulate a set of genes similar to that regulated by the deleted TF. Some of these paths (6 out of 88), such as Dal81→Pdr1→*HOR7* and Fkh2→Ace2→*ZWF1*, were already consistent with available data for both the source (e.g., Dal81 or Fkh2) and the intermediate (e.g., Pdr1 or Ace2) factor, because both TFs were already included among the 27 assayed ("Reinforcing" or "Feed-forward loop" in Fig. 5A). Other paths included intermediate factors for which the transcription profiles were implied but untested ("Indirect" in Fig. 5A), suggesting follow-up experiments to refine the model.

Swi6 is thought to bind DNA in a complex with either Swi4 or Mbp1 (*29*). To discriminate which Swi6 targets were Mbp1 dependent, we analyzed an *mbp1* knockout strain. The set of genes deletion-buffered by *mbp1Δ* was found to overlap with the *swi6Δ*-buffered set, including the genes *RNR1* and *DIN7* (Fig. 5C). Thus, *RNR1* and *DIN7* appear to depend on both Mbp1 and Swi6 for proper regulation.

Noting that loss of Rtg3 caused deletion buffering of many downstream genes through the path Rtg3→Ino4, we investigated whether loss of Rtg1, a binding partner of Rtg3 (*30*), would produce a similar outcome. Indeed, loss of Rtg1 buffered many genes that were bound by Ino4 (Fig. 5C). However, the sets of Rtg1 versus Rtg3 deletion-buffered genes did not strongly overlap. Thus, in response to MMS exposure, Rtg1 and Rtg3 appear to collaborate with Ino4 to regulate a battery of genes (whose products influence phospholipid metabolism and retrograde transport), but their functional roles are not interchangeable.

We have integrated TF binding profiles with genetic perturbations, mRNA expression, and protein interaction data to reveal direct and indirect interactions between TFs and MMS-responsive genes. The result is a highly interconnected physical map of regulatory pathways supported by both binding and deletion-buffering profiles. Some relations in this map are confirmed by previous studies, but most represent the basis for new hypotheses. As systems-level approaches continue to map the connectivity of large cellular systems, an important goal will be to make these maps even more integrative and to learn how to use them to predict the effects of different drugs, dosages, and genetic dispositions on pathway function.

References and Notes

1. E. C. Friedberg *et al.*, *DNA Repair and Mutagenesis* (American Society for Microbiology, Washington, DC, ed. 2, 2005).
2. J. Rouse, S. P. Jackson, *Science* **297**, 547 (2002).
3. A. P. Gasch *et al.*, *Mol. Biol. Cell* **12**, 2987 (2001).
4. S. A. Jelinsky, L. D. Samson, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 1486 (1999).
5. S. A. Jelinsky, P. Estep, G. M. Church, L. D. Samson, *Mol. Cell. Biol.* **20**, 8157 (2000).
6. T. J. Begley, A. S. Rosenbach, T. Ideker, L. D. Samson, *Mol. Cell* **16**, 117 (2004).
7. T. J. Begley, A. S. Rosenbach, T. Ideker, L. D. Samson, *Mol. Cancer Res.* **1**, 103 (2002).
8. M. Chang, M. Bellaoui, C. Boone, G. W. Brown, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 16934 (2002).
9. G. W. Birrell *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 8778 (2002).
10. T. Ideker *et al.*, *Science* **292**, 929 (2001).
11. T. I. Lee *et al.*, *Science* **298**, 799 (2002).
12. I. Simon *et al.*, *Cell* **106**, 697 (2001).
13. G. Habeler *et al.*, *Nucleic Acids Res.* **30**, 80 (2002).
14. H. W. Mewes *et al.*, *Nucleic Acids Res.* **30**, 31 (2002).
15. E. C. Friedberg, W. Siede, A. J. Cooper, in *The Molecular and Cellular Biology of the Yeast* Saccharomyces, J. R. Broach, J. Pringle, J. Elizabeth, Eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1991), vol. 1, pp. 147–192.
16. Materials and methods are available as supporting material on *Science* Online.
17. C. T. Harbison *et al.*, *Nature* **431**, 99 (2004).
18. J. Zhu, M. Q. Zhang, *Bioinformatics* **15**, 607 (1999).
19. E. Wingender *et al.*, *Nucleic Acids Res.* **29**, 281 (2001).
20. C. T. Workman, G. D. Stormo, *Pac. Symp. Biocomput.* **5**, 467 (2000).
21. T. C. Santiago, C. B. Mamoun, *J. Biol. Chem.* **278**, 38723 (2003).
22. E. A. Winzeler *et al.*, *Science* **285**, 901 (1999).
23. A. P. Gasch *et al.*, *Mol. Biol. Cell* **11**, 4241 (2000).
24. M. Huang, Z. Zhou, S. J. Elledge, *Cell* **94**, 595 (1998).
25. A. Chabes *et al.*, *Cell* **112**, 391 (2003).
26. V. R. Iyer *et al.*, *Nature* **409**, 533 (2001).
27. C. H. Yeang, T. Ideker, T. Jaakkola, *J. Comput. Biol.* **11**, 243 (2004).
28. I. Xenarios *et al.*, *Nucleic Acids Res.* **30**, 303 (2002).
29. C. Koch, T. Moll, M. Neuberg, H. Ahorn, K. Nasmyth, *Science* **261**, 1551 (1993).
30. Y. Jia, B. Rothermel, J. Thornton, R. A. Butow, *Mol. Cell. Biol.* **17**, 1110 (1997).

31. We acknowledge funding from the National Institute of Environmental Health Sciences (NIEHS), the National Cancer Institute (NCI), and the David and Lucille Packard Foundation. L.D.S. is the Ellison American Cancer Society Research Professor. We thank R. Young, D. Odom, T. I. Lee, M. Daly, and B. Ren for assistance with ChIP-chip analysis and J. Kadonaga, W. McGinnis, and G. Kassavetis for consultations on the binding motif predictions. Microarray data have been deposited in the ArrayExpress database (accession numbers E-TABM-92 and E-TABM-93).

Supporting Online Material

www.sciencemag.org/cgi/content/full/312/5776/1054/DC1
Materials and Methods
Figs. S1 to S8
Tables S1 to S9
References

1 November 2005; accepted 17 April 2006
10.1126/science.1122088

Lamin A–Dependent Nuclear Defects in Human Aging

Paola Scaffidi and Tom Misteli*

Mutations in the nuclear structural protein lamin A cause the premature aging syndrome Hutchinson-Gilford progeria (HGPS). Whether lamin A plays any role in normal aging is unknown. We show that the same molecular mechanism responsible for HGPS is active in healthy cells. Cell nuclei from old individuals acquire defects similar to those of HGPS patient cells, including changes in histone modifications and increased DNA damage. Age-related nuclear defects are caused by sporadic use, in healthy individuals, of the same cryptic splice site in lamin A whose constitutive activation causes HGPS. Inhibition of this splice site reverses the nuclear defects associated with aging. These observations implicate lamin A in physiological aging.

Mutations in the lamin A gene (*LMNA*) are responsible for the premature aging disease Hutchinson-Gilford progeria syndrome (*1*–*3*). The most prevalent HGPS mutation (heterozygous Gly^{608}→Gly^{608} with C changed to T) leads to a splicing defect and consequent generation of a truncated, dominant gain-of-function lamin A isoform (*2*, *3*). HGPS patient cells have various defects in nuclear structure and function (*4*–*6*). They are characterized by dysmorphic nuclear shape (*4*), increased DNA damage (*5*), and down-regulation of several nuclear proteins, including the heterochromatin protein HP1 and the LAP2 group of lamin A–associated proteins (*6*). Furthermore, HGPS cells have altered histone modification patterns, including reduced heterochromatin-specific trimethylation of Lys^9 on histone H3 (Tri-Me-K9H3) (*6*).

It has not been clear how HGPS relates to normal aging and whether lamin A plays any role in the physiological aging process (*7*, *8*). To address these questions, we determined whether HGPS-like nuclear defects occurred in cells from normally aged individuals. Multiple skin fibroblast cell lines from old (81 to

National Cancer Institute (NCI), NIH, Bethesda, MD 20892, USA.

*To whom correspondence should be addressed. E-mail: mistelit@mail.nih.gov

96 years) individuals consistently showed nuclear aberrations similar to those seen in HGPS cells (Fig. 1, A to D) (*6*). Whereas most cells from young individuals (3 to 11 years) showed robust staining for HP1, LAP2s, and Tri-Me-K9H3, a significant subpopulation of nuclei in cells from old individuals had reduced signals, similar to previous observations in HGPS cells (Fig. 1A) (*6*). Quantitative single-cell analysis demonstrated that the distributions of marker levels were significantly different in cells from young compared with old individuals (Fig. 1, B and C) ($P < 0.001$). In cells from old individuals, between 40 and 90% (average of 61 ± 23%) of cells showed a reduction of Tri-Me-K9H3 (*9*), whereas this fraction was 20 to 32% (average of 25 ± 3.8%) in cells from young individuals (Fig. 1, B to D). The percentage of cells with reduced HP1 and Tri-Me-K9H3 in cells from old donors was comparable to that found in HGPS patient cells, although a stronger overall reduction was observed in HGPS cells, with many cells showing nearly no signal (Fig. 1, B to C). Similar differences were found for LAP2s.

The nuclear defects in HGPS patient cells have previously been shown to accumulate during passage in culture (*4*). Cell lines from both young and old donors exhibited increased defects with prolonged passage (Fig. 1, E and F). However, defects accumulated much more rapidly in cells from old donors than in cells from young individuals (Fig. 1F and fig. S1). A further hallmark of HGPS cells is an increased amount of unrepaired DNA damage as indicated by the presence of foci containing phosphorylated histone H2AX (*5*). Similar to HGPS patient cells, the percentage of nuclei with phospho-H2AX foci was significantly higher in cells from old individuals (43 ± 6%) compared with young individuals (12 ± 3%; $P < 0.001$) (Fig. 1, G and H). The accumulation of H2AX foci in aging human cells is consistent with similar recent observations in baboons and mice (*10*, *11*). We conclude that cells from HGPS patients and normally aged individuals share several common nuclear defects.

The vast majority of HGPS cases are caused by the constitutive activation of a cryptic splice site in exon 11, which leads to aberrant removal of the 3′-terminal 150 nucleotides of this exon (*2*, *3*). The resulting Δ150 *LMNA* mRNA gives rise to a dominant gain-of-function lamin A isoform containing an internal deletion of 50 amino acids (Δ50 lamin A) (*2*, *3*). Because the cryptic splice site in the wild-type *LMNA* gene is highly homologous to a generic consensus splice donor site (Fig. 2A), we asked whether this cryptic splice site is also used by default in cells from healthy individuals. Reverse transcription polymerase chain reaction (RT-PCR) using primers in exon 9 and 12 amplified a minor truncated product in cells from healthy individuals of all ages (Fig. 2, B and C). DNA sequencing confirmed the identity of the product. Furthermore, all cell lines tested expressed the Δ150 *LMNA* isoform when detected with a pair of primers specific for the aberrant splice junction (Fig. 2, B and D). Transfection of a splicing reporter in multiple cell lines from both young and old donors confirmed the sporadic use of the cryptic splice site in the wild-type *LMNA* transcript (Fig. 2, F and G). Quantitative RT-PCR

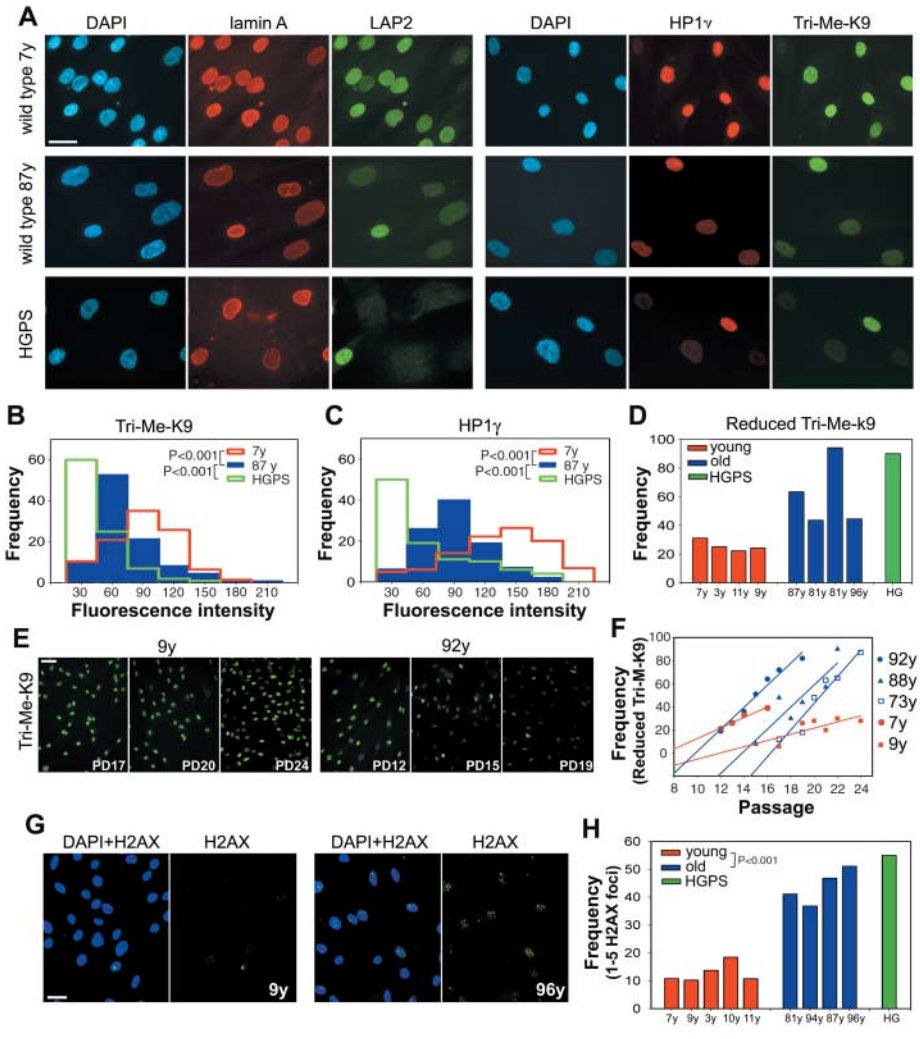


Fig. 1. Nuclear abnormalities in cells from old individuals. (**A**) Immunofluorescence microscopy on primary dermal fibroblasts from young (7 y) and old (87 y) healthy individuals and a HGPS patient. Scale bar, 10 μm. DAPI, 4′,6′-diamidino-2-phenylindole. Intensity distributions of the average fluorescent signal for (**B**) Tri-Me-K9H3 and (**C**) HP1γ in fibroblasts from healthy individuals of indicated age and a HGPS patient. (**D**) Quantitation of cells showing reduced amounts of Tri-Me-K9 in passage-matched population doublings (PD) 19 cell lines from healthy individuals of indicated age and a HGPS patient (*9*). (**E** and **F**) Reduction of Tri-Me-K9 H3 over cell passage. Best linear fits are shown. (**G** and **H**) Increased DNA damage in passage-matched (PD20) cell lines from healthy individuals of indicated age and a HGPS patient detected by antibody to H2AX. Scale bar, 20 μm. $N > 200$.

analysis using the splice junction primers demonstrated that use of the cryptic splice site was about 50-fold lower in cells from healthy individuals than in HGPS cells and did not increase with age (Fig. 2H). Use of the cryptic splice site is not a cell-culture artifact, because Δ150 *LMNA* mRNA is detectable in unprocessed human tissues (Fig. 2I). Similar to cultured cells, no significant difference in the amount of Δ150 *LMNA* mRNA was observed in tissues from either young or old individuals (Fig. 2I).

The occurrence of lamin A aberrant splicing in wild-type cells suggests that Δ50 lamin A is present in cells from healthy individuals. A truncated protein of the expected molecular weight was detected in wild-type cells by antibodies against lamin A and C (lamin A/C) upon extraction of soluble proteins and enrichment for membrane-associated proteins (Fig. 3A and fig. S2). Analysis of liver tissues confirmed the presence of Δ50 lamin A in old individuals (Fig. 3B). Consistent with the absence of any increase in Δ150 *LMNA* mRNA levels, the truncated protein did not significantly accumulate during aging (Fig. 3A). However, a striking change in lamin A/C localization was detected in cells from old donors (Fig. 3, C to E). Whereas in cells from young individuals a substantial fraction of lamin A/C was present throughout the nucleoplasm, this fraction was almost completely absent in cells from old donors, and the vast majority of lamin A accumulated at the nuclear rim (Fig. 3, C and D). This distribution was similar to that in HGPS cells, where the presence of the mutant lamin A protein leads to the accumulation of wild-type lamin A at the nuclear periphery (Fig. 3C) (*6*). Importantly, lack of nucleoplasmic lamin A/C strongly correlated with reduced amounts of Tri-Me-K9H3 and HP1 at the single-cell level (Fig. 3D). Thus, as in HGPS patients, cells from old individuals express the Δ150 *LMNA* mRNA, and lamin A is aberrantly localized in the cell nucleus.

To directly demonstrate that the production of the Δ50 lamin A protein by sporadic use of the cryptic splice site in *LMNA* exon 11 was responsible for the observed changes in nuclear architecture in cells from aged individuals, we inhibited the cryptic splice site in cells from old donors by using a previously characterized oligonucleotide-based approach (*6*). A morpholino oligonucleotide targeted to the cryptic splice site (Exo11) was introduced by repeated electroporation into cells from old individuals. We previously demonstrated that the Exo11 oligonucleotide efficiently inhibits the use of the cryptic splice site in vivo (*6*). Treatment of multiple cell lines from old individuals with Exo11 reversed their aberrant cellular phenotypes (Fig. 4). Fluorescent signals of HP1 and Tri-Me-K9H3 increased in cells treated with Exo11 compared with the same cells treated with a scrambled control oligonucleotide (Fig. 4A).

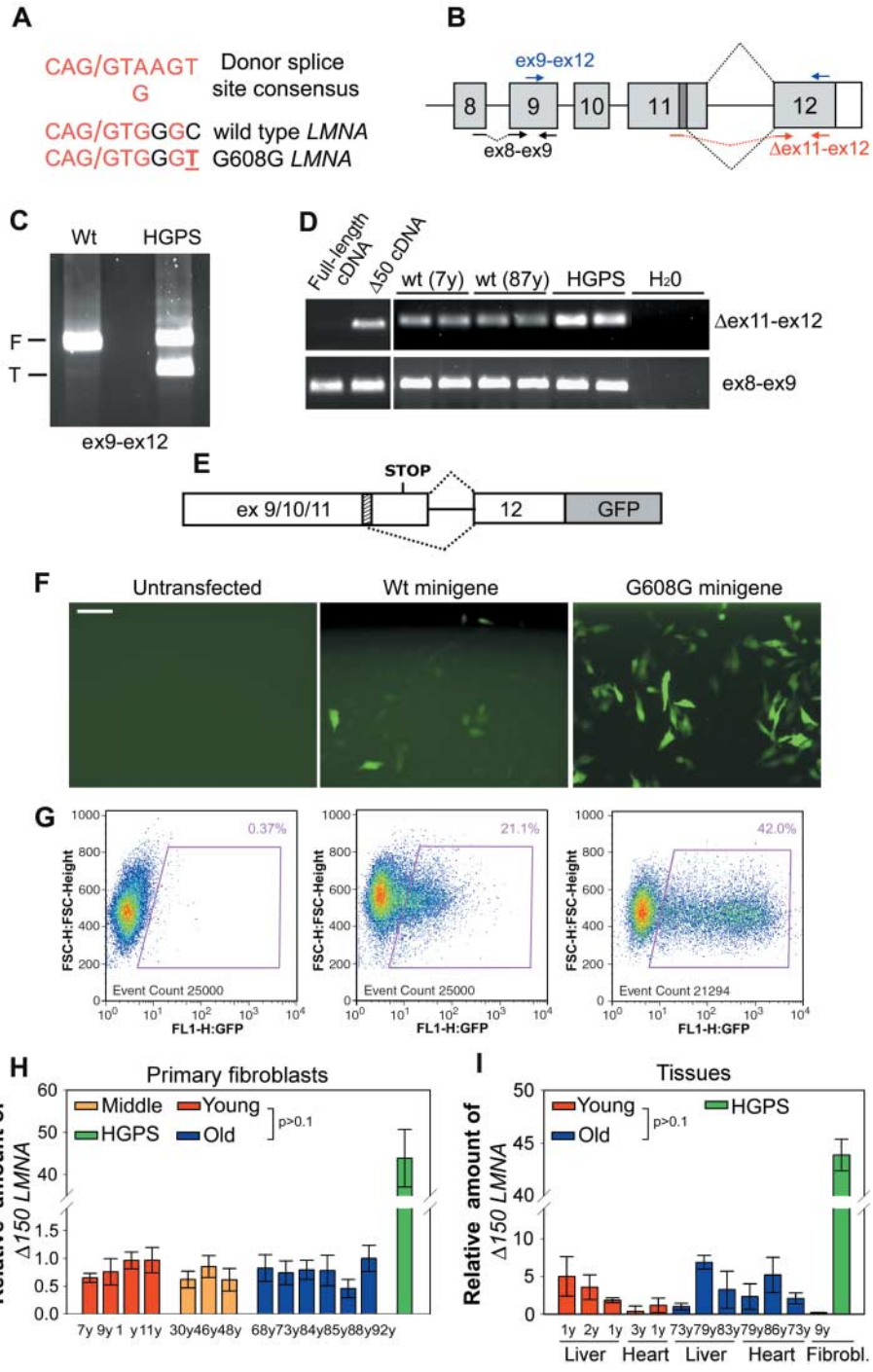


Fig. 2. Sporadic use of exon 11 *LMNA* cryptic splice site in cells from healthy individuals. (**A**) Comparison between the consensus donor splice sequence, the wild-type *LMNA* cryptic splice site in exon 11, and the mutated cryptic splice site in HGPS patients. Forward slash indicates the splice junction. (**B**) Schematic representation of splicing events (dotted black lines) involving *LMNA* exon 11 and exon 12. The cryptic splice site (dark box) and the position of different primer sets used for RT-PCR are indicated. RT-PCR analysis of wild-type and HGPS fibroblasts using (**C**) primers detecting both the full-length (F) and the truncated (T) *LMNA* isoforms or (**D**) primers specific for Δ150 *LMNA* (Δ ex11-ex12). Control ex8-ex9 primers detect all *LMNA* transcripts. Samples were analyzed in duplicate. Plasmids containing either full-length or truncated lamin A cDNA were used as controls for specificity of Δ150 *LMNA* detection. (**E**) Schematic representation of *LMNA* splicing reporter (*9*). The cryptic splice site (dashed box) is indicated. (**F**) Immunofluorescence microscopy and (**G**) fluorescence-activated cell sorting analysis of fibroblasts from a healthy individual transiently transfected with the indicated reporter minigene or untransfected. Scale bar, 60 μm. Quantitative RT-PCR analysis of (**H**) wild-type fibroblasts and (**I**) tissues from healthy individuals of indicated age and a HGPS patient. Δ150 *LMNA* RNA detected by using Δex11-ex12 primers relative to total *LMNA* RNA detected by using ex8-ex9 primers is shown.

Fig. 3. Accumulation of lamin A/C at the lamina in cells from old individuals. Western blot analysis of total protein extract and detergent insoluble fraction of fibroblasts from (**A**) healthy individuals of indicated age, a HGPS patient, and mouse fibroblasts and (**B**) from liver from healthy individuals probed with anti-lamin A/C. Δ50 lamin A is not detected in mouse fibroblasts. A fourth protein consistent with Δ10 lamin A (*15*) is also detected. Equal loading was verified by Coomassie staining of the blot. (**C** and **D**) Immunofluorescence confocal microscopy on passage-matched (PD21) fibroblasts. (D) Co-staining of fibroblasts from an old individual. Scale bar, 10 μm. Lamin A/C signal shown in intensity pseudocolors in upper right image. Arrowheads indicate nuclei with depleted nucleoplasmic lamin A/C and reduced amounts of HP1γ and Tri-Me-K9H3. (**E**) Quantitation of percentage of cells showing depletion of nucleoplasmic lamin A/C (average fluorescence intensity ratio nucleoplasm/lamina < 0.2) in passage-matched (PD21) cell lines from healthy individuals of indicated age and a HGPS patient.

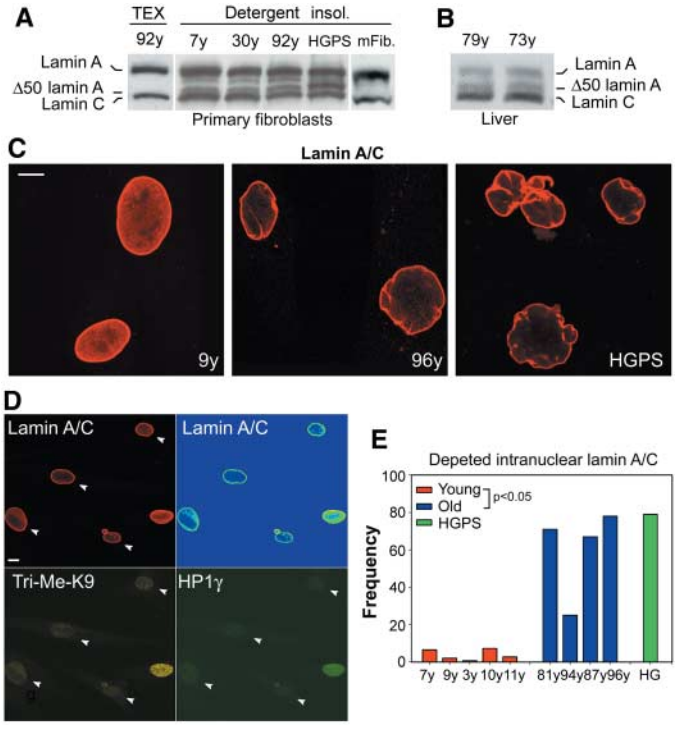
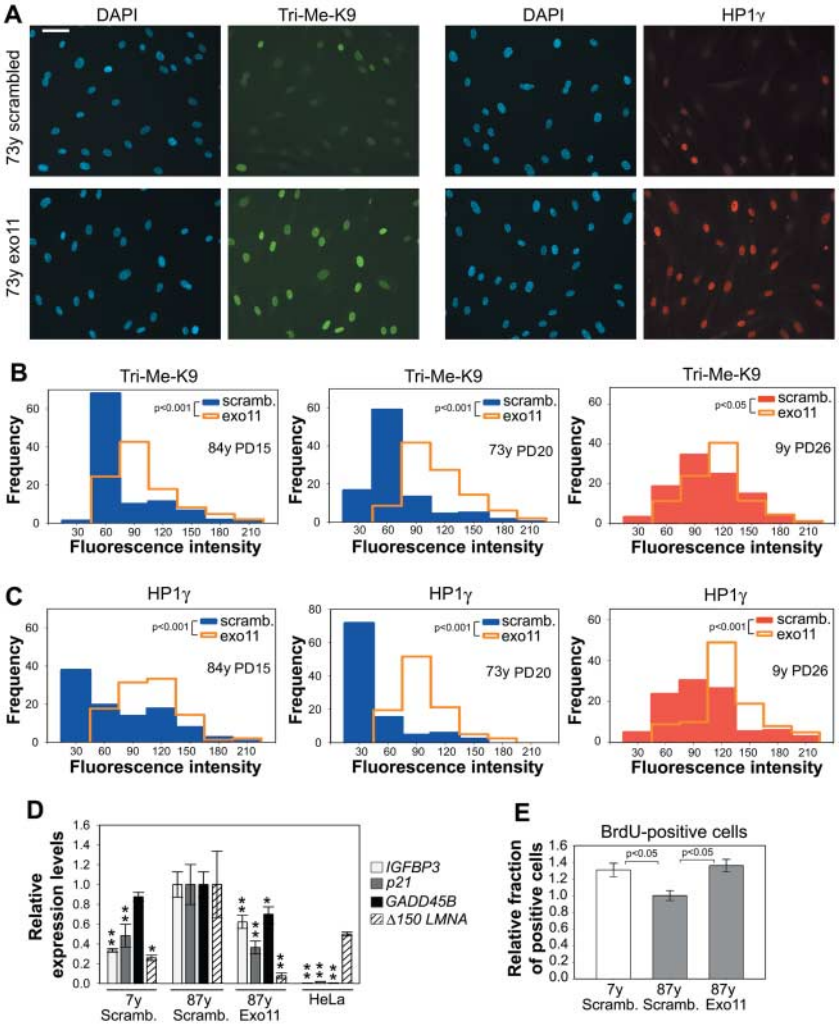


Fig. 4. Nuclear defects in cells from old individuals are caused by use of the *LMNA* exon 11 cryptic splice site. (**A**) Immunofluorescence microscopy on fibroblasts from an old healthy individual after treatment with either Exo11 oligonucleotide or a control scrambled oligonucleotide. Scale bar, 60 μm. (**B** and **C**) Quantitation of phenotypic rescue upon treatment with Exo11 oligonucleotide. (**D**) Quantitative RT-PCR analysis of HeLa cells and cell lines from young and old donors upon treatment with either Exo11 oligonucleotide or a control scrambled oligonucleotide. Values represent averages ± SD from a representative experiment. Statistical significances of the differences compared with mock-treated cells from the old donor are indicated. One asterisk indicates $P < 0.1$. Two asterisks indicate $P < 0.05$. (**E**) Quantitation of BrdU-positive cells in cell lines from young and old donors upon treatment with either Exo11 oligonucleotide or a control scrambled oligonucleotide. Values represent averages ± SD from three independent experiments.



Quantitation of the distribution of HP1 and Tri-Me-K9H3 confirmed the restoration of amounts similar to those seen in cells from young individuals at similar passages (Fig. 4, B and C) ($P < 0.001$). The reversibility of nuclear defects upon blocking of the cryptic splice site in *LMNA* directly demonstrates that lamin A is causal in generating the age-related nuclear defects. A smaller, but statistically significant, effect was also seen in cells from young individuals at later passages, further supporting a correlation between age and lamin A–dependent nuclear abnormalities (Fig. 4, B and C).

Organismal aging has been linked to activation of p53-dependent signaling pathways and initiation of the senescence program in a premature aging mouse model (*12*). To test whether lamin A plays a role in the activation of p53-dependent signaling pathways, we probed the status of p53 target genes after elimination of Δ50 lamin A from cells of old individuals using the Exo11 oligonucleotide. Inhibition of lamin A aberrant splicing resulted in significant down-regulation of *p21*, *IGFBP3*, and *GADD45B* compared with that of mock-treated cells (Fig. 4D). The status and response of p53 target genes were variable among cell lines. Consistent with the reduction in p53 activation, upon elimination of Δ50 lamin A from old cells, the fraction of 5-bromo-2′-deoxyuridine (BrdU)–positive proliferating cells increased by 30% and was similar to that in mock-treated young cells (Fig. 4E) ($P < 0.05$).

The sum of our observations implicates lamin A in physiological aging, showing that the same molecular mechanism responsible for the premature aging disease HGPS acts at a low level in healthy cells. The observation that HGPS and physiological aging share a common cellular and molecular basis strongly supports the notion that HGPS mimics at least some aspects of physiological aging and suggests that lamin A participates in the aging process in healthy individuals. Accelerated aging in HGPS might thus reflect an exaggerated lamin A–dependent mechanism, which normally contributes to physiological aging.

The fact that inhibition of the *LMNA* aberrant splicing event reverses the age-related defects in nuclear structure demonstrates that these abnormalities are caused by the Δ50 isoform of lamin A. Δ50 lamin A lacks an internal proteolytic processing site that is normally used as part of the lamin A maturation process but which is removed during the aberrant splicing event (*13*). The uncleaved pre–lamin A intermediate accumulates at the nuclear rim (*14*), where it is thought to exert its dominant effects (*12*, *14*). We show here that sporadic use of the cryptic splice site in exon 11 in healthy individuals leads to the production of the same pre–lamin A intermediate. Our observation that neither the amount of aberrantly spliced *LMNA* mRNA nor the amount of Δ50 lamin A increases during aging suggests that the nuclear defects are due to the prolonged presence of Δ50 lamin A in the nucleus. Aged cells might be more sensitized to the presence of the aberrant lamin A isoforms and less able to neutralize the negative effects of Δ50 lamin A, possibly due to the existence of a p53-dependent checkpoint, which senses structural abnormalities of the nuclear lamina and links those to the activation of the senescence program (*12*). Given our finding that several nuclear defects in aged cells are reversible upon inhibition of the aberrant splicing event in *LMNA*, it will be interesting to determine whether other cellular features of aging respond to such treatment and whether organismal aging can be modulated by interference with lamin A.

References and Notes

1. C. J. Hutchison, *Nat. Rev. Mol. Cell Biol.* **3**, 848 (2002).
2. A. De Sandre-Giovannoli *et al.*, *Science* **300**, 2055 (2003); published online 17 April 2003 (10.1126/science.1084125).
3. M. Eriksson *et al.*, *Nature* **423**, 293 (2003).
4. R. D. Goldman *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 8963 (2004).
5. B. Liu *et al.*, *Nat. Med.* **11**, 780 (2005).
6. P. Scaffidi, T. Misteli, *Nat. Med.* **11**, 440 (2005).
7. W. Andrew, *Adv. Gerontol. Res.* **18**, 87 (1964).
8. E. Haithcock *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 16690 (2005).
9. Materials and methods are available as supporting material on *Science* Online.
10. U. Herbig, M. Ferreira, L. Condel, D. Carey, J. M. Sedivy, *Science* **311**, 1257 (2006); published online 2 February 2006 (10.1126/science.1122446).
11. O. A. Sedelnikova *et al.*, *Nat. Cell Biol.* **6**, 168 (2004).
12. I. Varela *et al.*, *Nature* **437**, 564 (2005).
13. S. G. Young, L. G. Fong, S. Michaelis, *J. Lipid Res.* **46**, 2531 (2005).
14. S. H. Yang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 10291 (2005).
15. B. M. Machiels *et al.*, *J. Biol. Chem.* **271**, 9249 (1996).

16. We thank D. Donato for help with statistical analysis, C. Baker for design of specific primer pairs, M. Sinensky for advice on metabolic labeling of farnesylated lamin A, and K. Wilson and T. Jenuwein for providing reagents. Fluorescence imaging was performed at the NCI Fluorescence Imaging Facility. This research was supported by the Intramural Research Program of the NIH, NCI, Center for Cancer Research.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1127168/DC1
Materials and Methods
Figs. S1 and S2
References

8 March 2006; accepted 12 April 2006
Published online 27 April 2006;
10.1126/science.1127168
Include this information when citing this paper.

Biogeographic Evolution of Madagascar's Microendemic Biota

Lucienne Wilmé,[1] Steven M. Goodman,[2,3]* Jörg U. Ganzhorn[4]

The endemic species richness on Madagascar, relative to landmass area, is unparalleled in the world. Many organisms on the island have restricted geographical ranges. A comprehensive hypothesis explaining the evolution of this microendemism has yet to be developed. Using an analysis of watersheds in the context of Quaternary climatic shifts, we provide a new mechanistic model to explain the process of explosive speciation on the island. River catchments with sources at relatively low elevations were zones of isolation and hence led to the speciation of locally endemic taxa, whereas those at higher elevations were zones of retreat and dispersion and hence contain proportionately lower levels of microendemism. These results provide a framework for biogeographic and phylogeographic studies, as well as a basis for prioritizing conservation actions of the remaining natural forest habitats on the island.

Madagascar is renowned for its biodiversity and high levels of local endemism, particularly among forest species (*1*). Proportionate to land area, there is no other zone of the world with higher concentrations of biotic endemism across different taxonomic levels (*2*), and explaining the "origins of the modern terrestrial … fauna of Madagascar remain[s] one of the greatest unsolved mysteries of natural history" (*3*). At taxonomic ranks at and above the genus, the notable degree of endemism can be partially explained by the long isolation of Madagascar from Africa (more than 150 million years) and from India (less than 90 million years) (*4*). Infrequent nonsynchronized colonization by animals, and subsequent radiations, gave rise to a largely endemic biota (*5*–*7*). However, the processes that led to the evolution of many of these radiations within the extant fauna—most of which are forest dwelling and have high species-level turnover on very small geographic scales—have yet to be explained. Several attempts have been made to address these questions, but often at geographically local or taxonomically restricted levels (*8*). Neither adaptations to the present vegetation formations nor rivers as barriers provide a global framework for explaining the present biogeographic distribution of a considerable proportion of the island's fauna (*9*–*11*).

[1]Missouri Botanical Garden, Boîte Postale 3391, Antananarivo (101), Madagascar. [2]Field Museum of Natural History, 1400 South Lake Shore Drive, Chicago, IL 60605, USA. [3]WWF, Boîte Postale 738, Antananarivo (101), Madagascar. [4]Department of Animal Ecology and Conservation, University of Hamburg, 20146 Hamburg, Germany.

*To whom correspondence should be addressed. E-mail: sgoodman@fieldmuseum.org

We propose a new hypothesis to explain the evolutionary history of regional speciation in Madagascar's forest biota. Our hypothesis is based on a detailed database of the island's rivers and associated watersheds and an analysis of 35,400 geo-referenced records of modern land vertebrate species on the island. This mechanistic approach uses various parameters, different from previous models, to address the effects of paleoclimatic shifts on patterns of dispersal and vicariance at an intra-island level in recent geological time. Data on the distribution of modern Malagasy vertebrates were entered into the Noe4D program (*12*); data sources included verified specimens (housed in 34 museums across the world), scientific literature, and published field surveys (*13*). We used digital files to construct overlays, including topographical and hydrological layers, derived from Foiben-Taosarintanin'i Madagasikara maps (*13*).

Quaternary paleoclimatic variation has played an important role in the distribution and speciation of organisms on continental landmasses (*14*), and these same fluctuations have influenced Madagascar (*15*, *16*). During periods of glaciation, when the climate was cooler and drier, natural habitats at lower elevations experienced more-pronounced arid conditions than did zones at higher elevations (*17*). The precise sequence of cooler and drier periods is not well documented for Madagascar and cannot be easily extrapolated from other regions (*18*, *19*), but Quaternary shifts of ambient temperature on the island of more than 4°C have been postulated (*20*). Even with these climatic shifts, orographic precipitation allowed for the continuation of river systems and associated riverine forest habitat along hydrological systems that had sources in upper montane zones. We propose that these riverine habitats acted as buffers for the maintenance of more mesic local conditions and potential corridors for retreat toward higher altitudinal zones. However, critical for the model presented here, the influence of these climatic shifts was not equal across watersheds, and those with sources at relatively low elevations would have experienced more-notable ecological shifts, associated with aridification, and greater levels of habitat isolation than those occurring at higher elevations.

On Madagascar, there are three mountains at altitudes above 2000 m: one in the north (Tsaratanana), one in the center (Ankaratra), and one in the south (Andringitra). All of these zones have a network of rivers with their sources toward their respective summits (Fig. 1) and would have provided a direct means of retreat toward higher altitudinal zones during the vicissitudes of Quaternary or late Tertiary glaciation. We analyzed the biogeographic importance of rivers that had sources in different elevational zones, and their associated watersheds, to explain certain aspects of modern animal distributions. The divides of those rivers with their sources at upper elevations are considered here as "retreat-dispersion watersheds." During periods of glacial

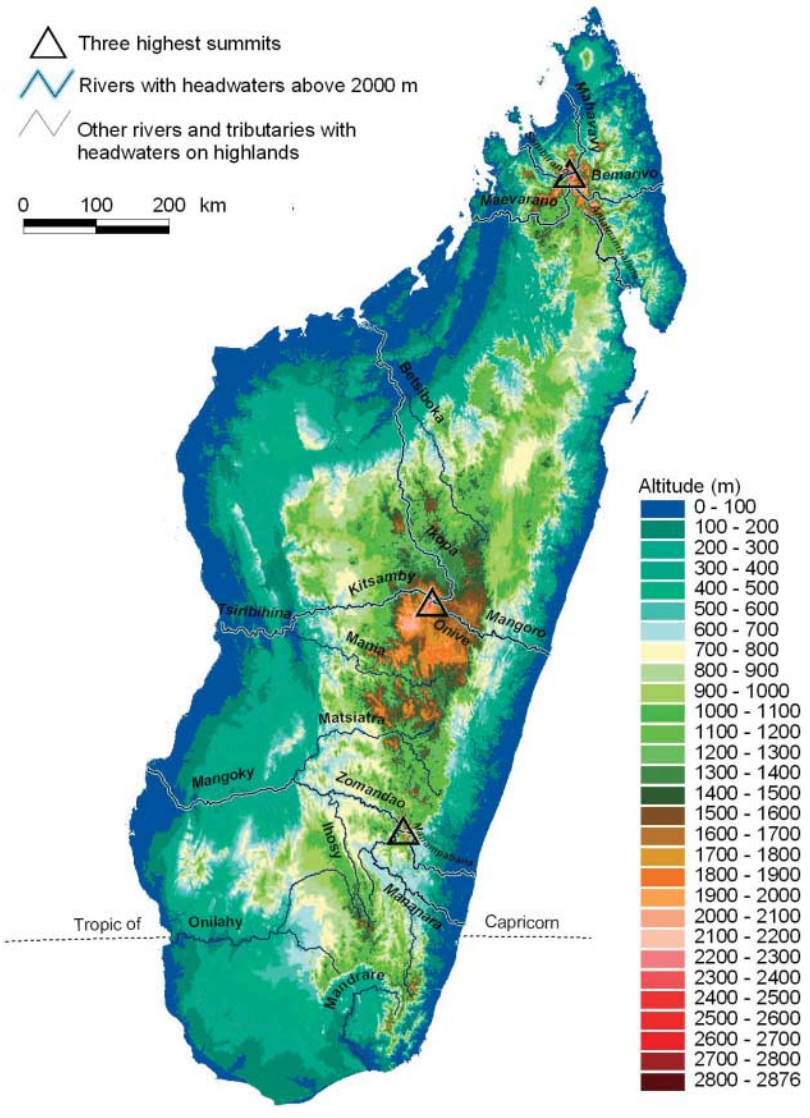


Fig. 1. Topography of Madagascar showing the three highest summits reaching altitudes above 2000 m.

minima, when the climatic conditions were notably warmer and more humid, as compared with glacial maxima with the inverse conditions, these montane riverine corridors would have provided the means for the expansion of geographical ranges for a variety of organisms (Fig. 2).

Entire populations of a given species occurring in the smaller watersheds, in between retreat-dispersion watersheds, would have been isolated through vicariant events, allowing for extensive allopatric speciation. These events would have been associated with, for example, temporal aspects of climatic shifts, mechanistic aspects of habitat differences, and isolation. The farther the watersheds are from the three highest summital zones on the island and the lower the headwaters are from their associated rivers, the more biotically isolated these areas would have been between retreat-dispersion watersheds.

This scenario predicts several centers of endemism in the more lowland and coastal portions of the island (Fig. 2). For example, most diurnal taxa of lemurs have recent distributions consistent with the hypothesized centers of endemism. Several new species of western nocturnal lemurs of the genus *Microcebus* have been described or resurrected from synonymy (*21*), and their distributions are also concordant with our centers-of-endemism hypothesis.

Madagascar is in the process of increasing the current protected areas system on the island from about 3% of the island's surface to 10% over the next 5 years (*22*). A simplified analysis, overlaying the centers of endemism and remaining forest cover on the current protected areas, reveals several areas that are in need of further action (Fig. 3), including some that do not contain a single protected area (*13*). Other centers of endemism are underrepresented, such as center 6 (south of Mangoky watershed), a diverse area including the Mikea Forest, the Mahafaly Plateau, and the southern spiny bush, and center 3 (between Mangoro and Mananara watersheds) with extensive remaining natural humid forests and the ecologically di-

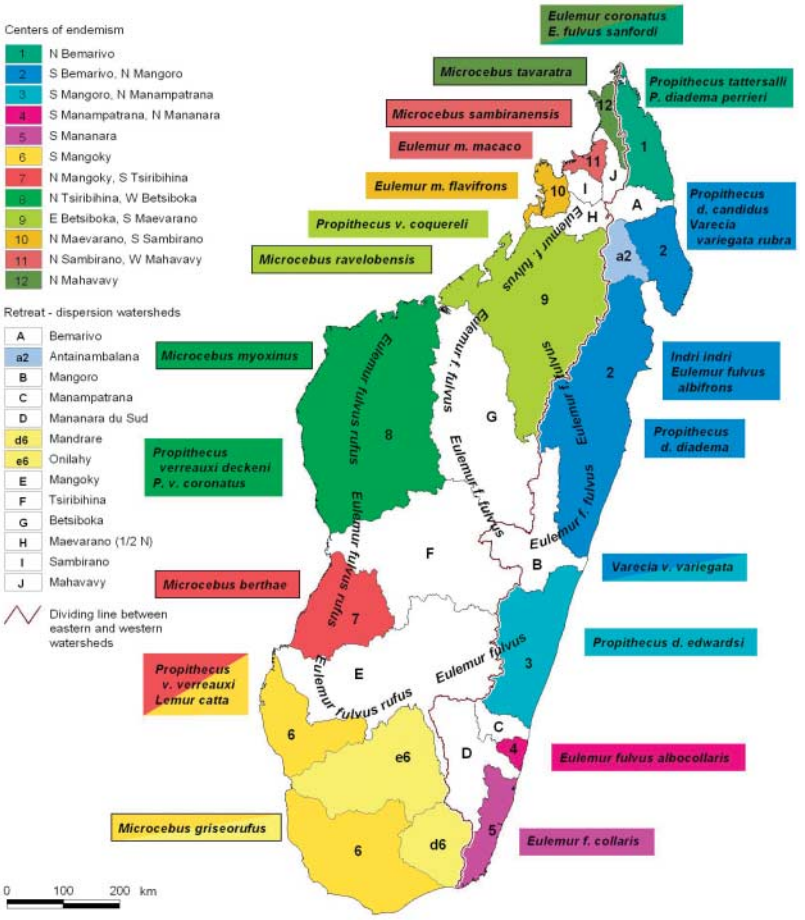


Fig. 2. The summital regions of the island are surrounded by highland areas and are the sources of major river drainages that serve as routes of retreat into refugia and subsequent dispersion (white; labeled with letters). Centers of endemism are assemblages of the smaller watersheds delimited between retreat-dispersion watersheds (colored; labeled with numbers) and with their sources at lower elevations. Recent distributions of diurnal lemurs (*Eulemur*, *Indri*, *Lemur*, *Propithecus*, and *Varecia*) show largely concordant patterns with the centers of endemism proposed here at the level of species or subspecies. In a few cases, the distribution of certain taxa crossover centers of endemism and these are designated with split coloration. Potential retreat-dispersion watersheds are labeled a2, d6, and e6. The complex of *E. f. fulvus*—*E. f. rufus* is broadly distributed across the retreat-dispersion watersheds and demonstrates broad regional genetic complexity. The newly recognized taxa in the genus *Microcebus* from the western portion of the island (outlined in black) show also consistent distribution with the centers of endemism presented here.

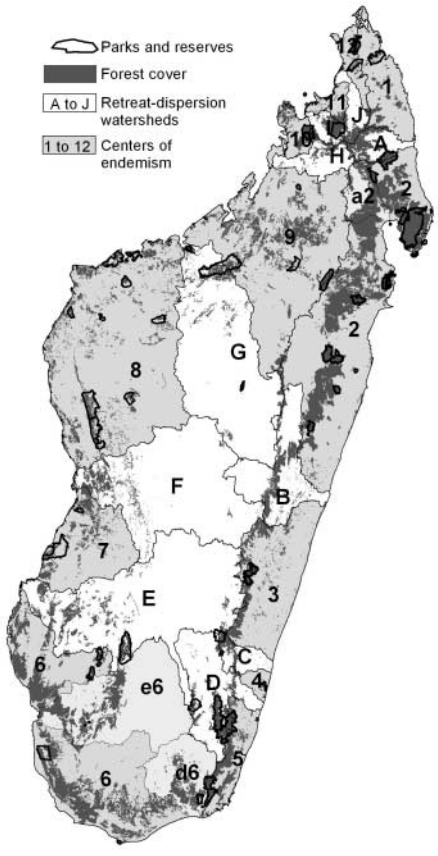


Fig. 3. Overlay between the centers of endemism, natural forest cover, and the existing protected areas system of Madagascar. The dark north-south line running the length of the island is the continental divide between eastern and western draining watersheds. Forest cover figures are from the 1994 "Inventaire Ecologique et Forestier National" assessment (*23*). See Fig. 2 for the key to the names of the centers of endemism.

verse forested zone between Fandriana and Marolambo and to the south of Ranomafana. The retreat-dispersal watersheds are also poorly represented, but in several cases these are regions with little remaining natural forest cover.

References and Notes

1. S. M. Goodman, J. P. Benstead, *Oryx* **39**, 73 (2005).
2. N. Myers, R. A. Mittermeier, C. G. Mittermeier, G. A. B. da Fonseca, J. Kent, *Nature* **403**, 853 (2000).
3. D. W. Krause, J. H. Hartman, N. A. Wells, in *Natural Change and Human Impact in Madagascar*, S. M. Goodman, B. D. Patterson, Eds. (Smithsonian Institution Press, Washington, DC, 1997), pp. 3–43.
4. M. Storey *et al.*, *Science* **267**, 852 (1995).
5. A. D. Yoder *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 6587 (2005).
6. F. Andreone, M. Vences, D. R. Vieites, F. Glaw, A. Meyer, *Mol. Phylog. Evol* **34**, 315 (2005).
7. C. Poux *et al.*, *Syst. Biol.* **54**, 719 (2005).
8. U. Thalmann, in *Diversité et Endemisme à Madagascar*, W. R. Lourenço, S. M. Goodman, Eds. (Mémoires de la Société de Biogéographie, Paris, 2000), pp. 191–202.
9. C. J. Raxworthy *et al.*, *Nature* **426**, 837 (2003).
10. R. D. Martin, *Philos. Trans. R. Soc. London Ser. B* **264**, 295 (1972).
11. S. M. Goodman, J. U. Ganzhorn, *J. Biogeogr.* **31**, 47 (2004).
12. O. Hertu, J.-M. Elouard, Nomenclature Œcologique et Environnementale (NOE), Logiciel Institut de Recherche pour le Développement: Legal deposition no. IDDN.FR.001.420001.R.C.1994.30200 at the Agence pour la Protection des Programmes (2005).
13. Materials and methods are available as supporting material on *Science* Online.
14. G. Hewitt, *Nature* **405**, 907 (2000).
15. H. Straka, in *Biogéographie de Madagascar*, W. R. Lourenço, Ed. [Editions Office de la Recherche Scientifique et Technique Outre-Mer (ORSTOM), Paris, 1995], pp. 37–47.
16. D. A. Burney, in *Natural Change and Human Impact in Madagascar*, S. M. Goodman, B. D. Patterson, Eds. (Smithsonian Institution Press, Washington, DC, 1997), pp. 75–89.
17. J. Haffer, *Science* **165**, 131 (1969).
18. J. Lynch-Stieglitz, *Science* **304**, 1919 (2004).
19. K. A. Hughen, T. I. Eglinton, L. Xu, M. Makou, *Science* **304**, 1955 (2004).
20. D. A. Burney *et al.*, *J. Hum. Evol.* **47**, 25 (2004).
21. R. M. Rasoloarison, S. M. Goodman, J. U. Ganzhorn, *Int. J. Primatol.* **21**, 963 (2000).
22. World Conservation Union (IUCN), *Benefits Beyond Boundaries. Proceedings of the 5th IUCN World Parks Congress* (IUCN, Gland, Switzerland, 2005), pp. 25–26.
23. J.-M. Dufils, in *The Natural History of Madagascar*, S. M. Goodman, J. P. Benstead, Eds. (Univ. Chicago Press, Chicago, 2003), pp. 88–96.

24. Funding was provided for the database by a grant from WWF. Support for fieldwork associated with this project has been received from Critical Ecosystem Partner Fund and Center for Biodiversity Conservation of Conservation International, Ellen Thorne Smith Fund of the Field Museum of Natural History, John D. and Catherine T. MacArthur Foundation, The National Geographic Society, The VolkswagenStiftung, and several different programs of WWF. For assistance and collaboration in our field research program, we are grateful to A. P. Raselimanana and other colleagues associated with the Ecology Training Program of WWF Madagascar Programme Office, to J.-M. Elouard and J.-M. Dufils for assistance with database and mapping, and to J.-M. Simonin for his early interest in the watersheds scenario.

Supporting Online Material

www.sciencemag.org/cgi/content/full/312/5776/1063/DC1
SOM Text
Fig. S1
Tables S1 to S3
References

18 November 2005; accepted 19 April 2006
10.1126/science.1122806

Gas Chromatograph/Mass Spectrometer

The QP-2010 Plus Gas Chromatograph/Mass Spectrometer (GC/MS) is suitable for rapid analysis of difficult samples and determination of trace components. Developed on the strengths of the model QP-2010, it features an extended mass range (from 1.5 to 1090 m/z) to allow analysis of high mass compounds and an extended ion source temperature range (from 100°C to 300°C) to increase flexibility and reduce maintenance. Dual turbo pumps provide increased sensitivity and column flow capacity, with twice the sensitivity for octafluoronapthalene and triple the sensitivity in electron impact and negative chemical ionization modes over the original. It is equipped with software allowing simultaneous scan/selected ion monitoring modes that enables gathering of detailed information on compounds of interest without sacrificing data across the entire mass spectrum.
Shimadzu Scientific Instruments For information 800-477-1227 www.ssi.shimadzu.com 593K

Transfection Reagent for Endothelial Cells

A powerful cationic polyethylenimine-derived transfection reagent optimized for the delivery of genes to fragile primary cells, such as human umbilical vein endothelial cells (HUVEC) and vascular endothelial cells of various origins, is available. As efficient as electroporation, jetPEI-HUVEC transfection reagent has demonstrated transfection efficiencies as high as 50% with little cytotoxicity.
Polyplus-transfection For information call 760-481-4918 www.polyplus-transfection.com

FT-IR and FT-NIR Spectrometers

The Spectrum 100 Fourier transform infrared (FT-IR) and 100N Fourier transform near infrared (FT-NIR) spectrometers offer high spectral quality and reproducible results with ease and efficiency. The series encompasses a broad array of features and advantages, including a new and improved attenuated total reflectance accessory, a new fiber optic NIR probe, and an enhanced version of PerkinElmer's Spectrum and AssureID software packages. This software provides a laboratory workflow-driven approach to testing materials for an easy-to-use measurement system that integrates seamlessly into operations in laboratories of any size.
Perkin-Elmer For information call 781-237-5100 www.perkinelmer.com

Mass Ion Chromatograms

The AccuScreen is a new method for high-resolution, accurate mass extracted ion chromatograms on the micrOTOF electron spray ionization–time of flight (ESI-TOF) mass spectrometers. AccuScreen yields superior liquid chromatography/mass spectrometry (LC/MS) performance in target compound analyses in complex matrices, such as body fluids, complex small molecule libraries, and quantitative proteomics. Applications include drug development studies; biomarker analysis; quantitative proteomics; and high-resolution screening for drugs of abuse, poisons, or toxic chemicals. The AccuScreen can generate accurate mass ion chromatograms with a tolerance down to ±0.002 Da. The benchtop micrOTOF series features outstanding sensitivity and stability, as well as mass accuracy of better than 3 ppm with internal calibration and mass resolution greater than 15,000 (FWHM).
Bruker Daltonics For information 978-663-3660 www.bdal.com

Nano Liquid Chromatography System

The NanoFrontier Nano Liquid Chromatography system is designed to meet the demanding challenges of today's research laboratories. The instrument features the patent-pending Dual Exchange Gradient System technology for gradient and retention time reproducibility at true nanoflow rates (50 nL/min). For complex proteomics applications, it can be configured for two-dimensional chromatography.
Hitachi High Technologies America
For information 800-548-9001 www.hitachi-hta.com

MALDI BioTyper System

The MALDI BioTyper system is for the identification and classification of microorganisms using protein "fingerprints" measured by matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) mass spectrometry. The characteristic protein expression patterns of microorganisms, such as bacteria, yeasts, and fungi, can be analyzed with the system. The reproducibility of the MALDI BioTyper is based on the measurement of high-abundance proteins, including many ribosomal proteins. Because ribosomal proteins are part of the cellular translational machinery, they are present in all living cells. As a result, the MALDI BioTyper protein fingerprints are not significantly influenced by variability in environmental or growth conditions. The BioTyper represents a complement to classical microbiological identification and classification techniques, and to modern polymerase chain reaction (PCR) systems. In contrast to other approaches, the BioTyper does not require any initial assessment like gram staining, oxidase test of unknown samples, choice of PCR primers, or use of selective growth media. Applications include taxonomical research on microorganisms, infectious disease research, and microorganism detection and identification in environmental analysis, food safety, and water quality.
Bruker Daltonics For information 978-663-3660 www.bdal.com

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and government organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier by visiting www.science.labvelocity.com on the Web, where you can request that the information be sent to you by e-mail, fax, mail, or telephone.